PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 5 :
C07K 7/10, A61K 39/21
C12N 15/48, C12P 21/02, 21/08
G01N 33/569

A1

(11) International Publication Number: WO 90/03984

(43) International Publication Date: 19 April 1990 (19.04.90)

(21) International Application Number: PCT/US89/04302

(22) International Filing Date: 29 September 1989 (29.09.89)

(30) Priority data:

252,949	3 October 1988 (03.10.88)	US
359,543	1 June 1989 (01.06.89)	US
407,663	19 September 1989 (19.09.89)	US

(71) Applicant: REPLIGEN CORPORATION [US/US]; One Kendall Square, Building 700, Cambridge, MA 02139 (US).

(72) Inventors: RUSCHE, James, R. ; 18 Brigham Road, Framingham, MA 01701 (US). PUTNEY, Scott, D. ; 5 Epping Street, Arlington, MA 02174 (US). JAVAHERIAN, Kashayar ; 27 Webster Road, Lexington, MA 02173 (US). FARLEY, John ; 507 Reeves Road, Pittsford, NY 14534 (US). GRIMAILA, Raymond ; 311 Washington Street, Somerville, MA 02134 (US). LYNN, Debra ; 1 Watermill Place, Unit 328, Arlington, MA 02174 (US). PETRO-BREYER, Joan ; Institut Pasteur, Department of Biotechnology, 25, rue du Docteur-Roux, F-75724 Paris Cédex 15 (FR). O'KEEFFE, Thomas ; 14 Karen Road, Windham, NH 03089 (US). LAROSA, Gregory ; 85 Strathmore Road, Apt. 33, Brookline, MA 02146 (US). PROFY, Albert, T. ; 28 Essex Street, Cambridge, MA 02139 (US).

(74) Agents: SALIWANCHIK, Roman et al.; Saliwanchik & Saliwanchik, 2421 N.W. 41st Street, Gainesville, FL 32606 (US).

(81) Designated States: AT (European patent), AU, BE (European patent), CH (European patent), DE (European patent), DK, FI, FR (European patent), GB (European patent), IT (European patent), JP, KR, LU (European patent), NL (European patent), NO, SE (European patent).

Published
With international search report.

(54) Title: HIV PROTEINS AND PEPTIDES USEFUL IN THE DIAGNOSIS, PROPHYLAXIS OR THERAPY OF AIDS

(57) Abstract

The subject invention concerns the identification of a portion of the HIV envelope protein called the principal neutralizing domain. Polypeptides comprising this domain have the capability of raising, and/or binding with, neutralizing antibodies. The invention further concerns novel HIV polypeptides which can be used in the diagnosis, prophylaxis, or therapy of AIDS. These polypeptides can be prepared by known chemical synthetic procedures, or by recombinant DNA means. The polypeptides pertain to the gp 120 subunit, amino acids 298-320, including the sequence ....gly-pro-gly...... and variants thereof. Multiepitope polypeptides comprising analogues of this peptide epitope from different HIV variants are referred to.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	ES	Spain	MG	Madagascar
AU	Australia	FI	Finland	ML	Mali
BB	Barbados	FR	France	MR	Mauritania
BE	Belgium	GA	Gabon	MW	Malawi
BF	Burkina Fasso	GB	United Kingdom	NL	Netherlands
BG	Bulgaria	HU	Hungary	NO	Norway
BJ	Benin	IT	Italy	RO	Romania
BR	Brazil	JP	Japan	SD	Sudan
CA	Canada	KP	Democratic People's Republic of Korea	SE	Sweden
CF	Central African Republic			SN	Senegal
CG	Congo	KR	Republic of Korea	SU	Soviet Union
CH	Switzerland	LI	Liechtenstein	TD	Chad
CM	Cameroon	LK	Sri Lanka	TG	Togo
DE	Germany, Federal Republic of	LU	Luxembourg	US	United States of America
DK	Denmark	MC	Monaco		

DESCRIPTION

HIV PROTEINS AND PEPTIDES USEFUL IN THE DIAGNOSIS, PROPHYLAXIS OR THERAPY OF AIDS

Cross-Reference to a Related Application

This is a continuation-in-part of our co-pending application Serial No. 359,543, filed on June 1, 1989, which is a continuation-in-part of our co-pending application Serial No. 252,949, filed on October 3, 1988, which is a continuation-in-part of our co-pending application Serial No. 090,080, filed on August 27, 1987.

Background of the Invention

Human immunodeficiency virus (HIV), human T-cell lymphotropic virus III (HTLV-III), lymphadenopathy-associated virus (LAV), or AIDS-associated retrovirus (ARV) has been identified as the cause of acquired immune deficiency syndrome (AIDS) (Popovic, M., M.G. Sarngadharan, E. Read, and R.C. Gallo [1984] Science 224:497-500). The virus displays tropism for the OKT4^{+} lymphocyte subset (Klatzmann, D., F. Barre-Sinoussi, M.T. Nugeyre, C. Dauget, E. Vilmer, C. Griscelli, F. Brun-Vezinet, C. Rouzioux, J.D. Gluckman, J.C. Chermann, and L. Montagnier [1984] Science 225:59-63). Antibodies against HIV proteins in the sera of most AIDS and AIDS related complex (ARC) patients, and in asymptomatic people infected with the virus (Sarngadharan, M.G., M. Popovic, L. Bruch, J. Schupbach, and R.C. Gallo [1984] Science 224:506-508) have made possible the development of immunologically based tests that detect antibodies to these antigens. These tests are used to limit the spread of HIV through blood transfusion by identifying blood samples of people infected with the virus. Diagnostic tests currently available commercially use the proteins of inactivated virus and antigens.

In addition to allowing new approaches for diagnosis, recombinant DNA holds great promise for the development of vaccines against both bacteria and viruses (Wilson, T. [1984] Bio/Technology 2:29-39). The most widely employed organisms to express recombinant vaccines have been E. coli, S. cerevisiae and cultured mammalian cells. For example, subunit vaccines against foot and mouth disease (Kleid, D.G., D. Yansura, B. Small, D. Dowbenko, D.M. Moore, M.J. Brubman, P.D. McKercher, D.O. Morgan, B.H. Robertson, and H.L. Bachrach [1981] Science 214:1125-1129) and malaria (Young, J.F., W.T. Hockmeyer, M. Gross, W. Ripley-Ballou, R.A. Wirtz, J.H. Trosper, R.L. Beaudoin, M.R. Hollingdale, L.M. Miller, C.L. Diggs, and M. Rosenberg [1985] Science 228:958-962) have been synthesized in E. coli. Other examples are hepatitis B surface antigen produced in yeast (McAleer, W.J., E.B. Buynak, R.Z. Maigetter, D.E. Wampler, W.J. Miller, and M.R. Hilleman [1984] Nature 307:178-180) and a herpes vaccine produced in mammalian cells (Berman, P.W., T. Gregory, D. Chase, and L.A. Lasky [1984] Science 227:1490-1492).

The entire HIV envelope or portions thereof have been used to immunize animals. The terms "protein," "peptide," and "polypeptide" have been used interchangeably in this application to refer to chemical compounds having more than one amino acid. The term "compound" as used here refers to chemical compounds in general. Thus, "compound" includes proteins, peptides, and polypeptides. Also included under the category of "compound" are fusion compounds where polypeptides are combined with non-polypeptide moieties. As used in the present application, the term "naturally occurring HIV envelope protein" refers to the proteins gp160, gp120, and gp41 only. As used in the present application, HIV refers to any HIV virus, including HIV-1 and HIV-2.

Both the native gp120 (Robey et al. [1986] Proc. Natl. Acad. Sci. 83:7023-7027; Matthews et al. [1986] Proc. Natl. Acad. Sci. 83:9709-9713) and recombinant proteins (Laskey et al. [1986] Science 233:209-212; Putney et al. [1986] Science 234:1392-1395) elicit antibodies that can neutralize HIV in cell culture. However, all of these immunogens elicit antibodies that neutralize only the viral variant from which the subunit was derived. Therefore, a novel vaccine capable of protecting against multiple viral variants would be advantageous and unique.

HIV is known to undergo amino acid sequence variation, particularly in the envelope gene (Starcich, B.R. [1986] Cell 45:637-648; Hahn, B.H. et al. [1986] Science 232:1548-1553). Over 100 variants have been analyzed by molecular cloning and restriction enzyme recognition analysis, and several of these have been analyzed by nucleotide sequencing. Some of these are the HIV isolates known as RF (Popovic, M. et al. [1984] Science 224:497-500), WMJ-1 (Hahn, B.H. et al. [1986] Science 232:1548-1553), LAV (Wain-Hobson, S. et al. [1985] Cell 40:9-17), and ARV-2 (Sanchez-Pescador, R. et al. [1985] Science 227:484-492). One aspect of this invention is defining the portion of HIV that comprises the principal neutralizing domain. The principal neutralizing domain is located between the cysteine residues at amino acids 296 and 331 of the HIV envelope. The numbering of amino acids follows the published sequence of HIV-III$_B$ (Ratner, L. et al. [1985] Nature 313:277-284). This domain is known to be hypervariable but retains the type-specific antigenic and immunogenic properties related to virus neutralization.

A further aspect of the subject invention is the discovery of highly conserved amino acids within the principal neutralizing domain. Although certain sequences from this region have been published (see, for example, Southwest Foundation for Biological Research, published PCT application, Publication No. WO 87/02775; Genetic Systems Corporation, Published United Kingdom Application No. GB 2196634 A; Stichting Centraal Diergeneeskundig Instituut, Published EPO Application No. 0 311 219), the presence of the conserved regions described here have never before been described.

Diagnostic kits or therapeutic agents using viral proteins isolated from virus infected cells or recombinant proteins would contain epitopes specific to the viral variant from which they were isolated. Reagents containing proteins from multiple variants would have the utility of being more broadly reactive due to containing a greater diversity of epitopes. This would be advantageous in the screening of serum from patients or therapeutic treatment of patients.

Synthetic peptides can be advantageous as the active ingredient in a vaccine, therapeutic agent or diagnostic reagent due to the ease of manufacture and ability to modify their structure and mode of presentation.

Synthetic peptides have been used successfully in vaccination against foot and mouth disease virus (Bittle et al. [1982] Nature 298:30-33); poliovirus (Emini et al. [1983] Nature 305:699); hepatitis B (Gerin et al. [1983] Proc. Natl. Acad. Sci. 80:2365-2369); and influenza (Shapira et al. [1984] Proc. Natl. Acad. Sci. 81:2461-2465).

There is a real need at this time to develop a vaccine for AIDS. Such a vaccine, advantageously, would be effective to immunize a host against the variant AIDS viruses.

Brief Summary of the Invention

The subject invention defines the location of the HIV principal neutralizing domain and discloses methods to utilize this segment of the HIV envelope protein for developing diagnostic, therapeutic, and prophylactic reagents. More specifically, the HIV principal neutralizing domain is located between cysteine residues 296 and 331 of the HIV envelope protein. The location of this domain is shown in Table 1. Although the specific amino acid sequence of the principal neutralizing domain is known to be highly variable between variants, we have found that peptides from this domain are invariably capable of raising, and/or binding with, neutralizing antibodies. This unexpected discovery provides a basis for designing compositions and strategies for the prevention, diagnosis, and treatment of AIDS.

The discovery of the principal neutralizing domain (also known as the "loop") resulted from extensive research involving a multitude of HIV envelope proteins and peptides from many HIV variants. Proteins and peptides capable of raising, and/or binding with, neutralizing antibodies are disclosed here. These novel HIV proteins and peptides, or their equivalents, can be used in the diagnosis, prophylaxis, and/or therapy of AIDS. Further, the peptides can be used as immunogens or screening reagents to generate or identify polyclonal and monoclonal antibodies that would be useful in prophylaxis, diagnosis and therapy of HIV infection.

A further aspect of the invention is the discovery of highly conserved regions within the principal neutralizing domain. This discovery was quite unexpected because of the known variability of the amino acids within this segment of the HIV envelope protein.

The proteins and peptides of the invention are identified herein by both their amino acid sequences and the DNA encoding them. Thus, they can be prepared by known chemical synthetic procedures, or by recombinant DNA means.

These peptides, or peptides having the antigenic or immunogenic properties of these peptides, can be used, advantageously, in a vaccine, e.g., a cocktail of peptides, to elicit broad neutralizing antibodies in the host. Further, these peptides can be used sequentially, e.g., immunizing initially with a peptide equivalent to the principal neutralizing domain of an HIV variety followed by immunization

with one or more of the above peptides. Polyclonal or monoclonal antibodies that bind to these peptides would be advantageous in prophylaxis or therapy against HIV, the causative agent of AIDS.

Brief Description of the Drawings

Figure 1 shows commonly occurring sequences of the principal neutralizing domain.

Figure 2 is a schematic for multi-epitope gene construction.

Figure 3 depicts the steps in the construction of a specific multi-epitope gene.

Figure 4 shows the sequences of four synthesized single-stranded oligomers for construction of a multi-epitope gene.

Detailed Disclosure of the Invention

Described here is a segment of the HIV envelope protein which raises, and/or binds with, neutralizing antibodies. This unique and highly unexpected property has been observed in each HIV variant that has been examined. The segment of interest has been named the "principal neutralizing domain." The principal neutralizing domain is bounded by cysteine residues which occur at positions 296 and 331. It should be noted that these same cysteine residues have been described as beginning at 302, rather than 296 (Rusche, J.R. et al. [1988] Proc. Natl. Acad. Sci. USA 85(15):3198-3202). Because the cysteine residues are linked through disulfide bonds, the segment between the residues tends to form a loop. Therefore, the principal neutralizing domain is also referred to as the "loop."

The segment of the protein envelope identified here as the principal neutralizing domain is known to be highly variable between HIV variants. Thus it is surprising that, for each variant, this segment is capable of eliciting, and/or binding with, neutralizing antibodies.

The principal neutralizing domain identified here is a small segment of the HIV envelope protein. This small segment may be combined with additional amino acids, if desired, for a specific purpose. All such proteins are claimed here except where such proteins constitute a naturally occurring HIV envelope protein. As used here, the term "naturally occurring envelope protein" refers only to gp160, gp120, and gp41.

Listed in Table 1 are sequences of the principal neutralizing domain for some of the variants tested. Table 9 contains a complete list of the principal neutralizing domains.

Amino acids may be referred to using either a three-letter or one-letter abbreviation system. The following is a list of the common amino acids and their abbreviations:

Amino acid	Three-letter symbol	One-letter symbol
Alanine	Ala	A
Arginine	Arg	R
Asparagine	Asn	N
Aspartic acid	Asp	D
Asn and/or Asp	Asx	B
Cysteine	Cys	C
Glutamine	Gln	Q
Glutamic acid	Glu	E
Gln and/or Glu	Glx	Z
Glycine	Gly	G
Histidine	His	H
Isoleucine	Ile	I
Leucine	Leu	L
Lysine	Lys	K
Methionine	Met	M
Phenylalanine	Phe	F
Proline	Pro	P
Serine	Ser	S
Threonine	Thr	T
Tryptophan	Trp	W
Tyrosine	Tyr	Y
Valine	Val	V

The following is a list of proteins and peptides which comprise principal neutralizing domains or segments thereof.

A. Recombinant Proteins Comprising a Principal Neutralizing Domain

1. HIV 10 Kd fusion protein denoted Sub 1. The amino acid sequence of the HIV portion of Sub 1 is shown in Table 2 and the DNA sequence of the HIV portion of Sub 1 in Table 2A. The amino acid sequence of Sub 1 is shown in Table 2B and the DNA sequence in Table 2C.

2. HIV 18 Kd fusion protein denoted Sub 2. The amino acid sequence of the HIV portion of Sub 2 is shown in Table 3 and the DNA sequence of the HIV portion of Sub 2 in Table 3A. The entire amino acid sequence of Sub 2 is shown in Table 3B and the entire DNA sequence in Table 3C.
3. HIV 27 Kd fusion protein denoted $PB1_{RF}$. The amino acid sequence of the HIV portion of $PB1_{RF}$ is listed in Table 4 and the DNA sequence of the HIV portion of $PB1_{RF}$ is listed in Table 4A. The entire amino acid sequence and DNA sequence of $PB1_{RF}$ are in Tables 4B and 4C, respectively.
4. HIV 28 Kd fusion protein denoted $PB1_{MN}$. The amino acid sequence of the HIV portion of $PB1_{MN}$ is shown in Table 5 and the DNA sequence of the HIV portion of $PB1_{MN}$ is shown in Table 5A. The entire amino acid sequence and DNA sequence of $PB1_{MN}$ are shown in Tables 5B and 5C, respectively.
5. HIV 26 Kd fusion protein denoted $PB1_{SC}$. The amino acid sequence of the HIV portion of $PB1_{SC}$ is listed in Table 6 and the DNA sequence of the HIV portion of $PB1_{SC}$ is shown in Table 6A. The entire amino acid sequence and DNA sequence of $PB1_{SC}$ are shown in Tables 6B and 6C, respectively.
6. HIV 26 Kd fusion protein denoted $PB1_{WMJ2}$. The amino acid sequence of the HIV portion of $PB1_{WMJ2}$ is listed in Table 7 and the DNA sequence of the HIV portion of $PB1_{WMJ2}$ is shown in Table 7A. The entire amino acid sequence and DNA sequence of $PB1_{WMJ2}$ are shown in Tables 7B and 7C, respectively.

B. Synthetic Peptides Comprising Segments of the Principal Neutralizing Domain From HIV Variants

The amino acid cysteine in parentheses is added for the purpose of crosslinking to carrier proteins. Also, where the peptides have cysteines at or near both ends, these cysteines can form a disulfide bond, thus giving the peptides a loop-like configuration. For any of these peptides which do not have cysteines at or near both ends, cysteines may be added if a loop-like configuration is desired. The loop configuration can be utilized to enhance the immunogenic properties of the peptides. Other amino acids in parentheses are immunologically silent spacers.

<u>Peptide 135</u> (from isolate HIV-III_B):

Asn	Asn	Thr	Arg	Lys	Ser	Ile	Arg	Ile	Gln	Arg	Gly
Pro	Gly	Arg	Ala	Phe	Val	Thr	Ile	Gly	Lys	Ile	Gly
(Cys)											

Peptide 136 (from isolate HIV-IIIB):

Leu Asn Gln Ser Val Glu Ile Asn Cys Thr Arg Pro Asn
Asn Asn Thr Arg Lys Ser Ile Arg Ile Gln Arg Gly Pro
Gly Arg Ala Phe Val Thr Ile Gly Lys Ile Gly Asn Met

Peptide 139 (from isolate HIV-RF):

Asn Asn Thr Arg Lys Ser Ile Thr Lys Gly Pro Gly
Arg Val Ile Tyr Ala Thr Gly Gln Ile Ile Gly (Cys)

Peptide 141 (from isolate HIV-WMJ2):

Asn Asn Val Arg Arg Ser Leu Ser Ile Gly Pro Gly
Arg Ala Phe Arg Thr Arg Glu Ile Ile Gly (Cys)

Peptide 142 (from isolate HIV-MN):

Tyr Asn Lys Arg Lys Arg Ile His Ile Gly Pro Gly
Arg Ala Phe Tyr Thr Thr Lys Asn Ile Ile Gly
(Cys)

Peptide 143 (from isolate HIV-SC):

Asn Asn Thr Thr Arg Ser Ile His Ile Gly Pro Gly
Arg Ala Phe Tyr Ala Thr Gly Asp Ile Ile Gly
(Cys)

Peptide 131 (from isolate HIV-IIIB):

(Tyr) Cys Thr Arg Pro Asn Asn Asn Thr Arg Lys Ser Ile
Arg Ile Gln Arg Gly

Peptide 132 (from isolate HIV-IIIB):

Pro Gly Arg Ala Phe Val Thr Ile Gly Lys Ile Gly Asn
Met Arg Gln Ala His Cys (Tyr)

Peptide 134 (from isolate HIV-IIIB):

Glu Arg Val Ala Asp Leu Asn Gln Ser Val Glu Ile Asn
Cys Thr Arg Pro Asn Asn Asn Thr Arg Lys Ser Ile

Peptide 339 (from isolate HIV-RF):

Ile Thr Lys Gly Pro Gly Arg Val Ile Tyr (Cys)

RP341 (from isolate HIV-WMJ2):

Leu Ser Ile Gly Pro Gly Arg Ala Phe Arg (Cys)

RP343 (from isolate HIV-SC):

Ile His Ile Gly Pro Gly Arg Ala Phe Tyr (Cys)

RP60 (from isolate HIV-IIIB):

Ile Asn Cys Thr Arg Pro Asn Asn Asn Thr Arg Lys Ser Ile

RP335 (from isolate HIV-IIIB):

Ile Gln Arg Gly Pro Gly Arg Ala Phe (Cys)

RP337 (from isolate HIV-IIIB):

Lys Ser Ile Arg Ile Gln Arg Gly Pro Gly Arg Ala Phe (Cys)

RP77 (from isolate HIV-IIIB):

Gly Pro Gly Arg Ala Phe

RP83 (from isolate HIV-WMJ1):

His Ile His Ile Gly Pro Gly Arg Ala Phe Tyr Thr Gly (Cys)

RP79 (from isolate HIV-IIIB):

Gln Arg Gly Pro Gly Arg Ala Phe (Cys)

RP57:

Ile Asn Cys Thr Arg Pro Ala His Cys Asn Ile Ser

RP55:

Ala His Cys Asn Ile Ser

RP75A:

(Ala Ala Ala Ala Ala Ala) Gly Pro Gly Arg (Ala Ala Ala Ala Ala Cys)

RP56:

Ile Asn Cys Thr Arg Pro

RP59:

Ile Gly Asp Ile Arg Gln Ala His Cys Asn Ile Ser

RP342 (from isolate HIV-MN):

Ile His Ile Gly Pro Gly Arg Ala Phe Tyr Thr (Cys)

RP96 (HIV-MN related):

(Cys) Gly Ile His Ile Gly Pro Gly Arg Ala Phe Tyr Thr (Cys)

RP97 (HIV-MN related):

(Ser Gly Gly) Ile His Ile Gly Pro Gly Arg Ala Phe Tyr (Gly Gly Ser Cys)

RP98 (HIV-MN related):

(Cys Ser Gly Gly) Ile His Ile Gly Pro Gly Arg Ala Phe Tyr (Gly Gly Ser Cys)

RP99 (HIV-MN related):

(Cys Ser Gly Gly) Ile His Ile Gly Pro Gly Arg Ala Phe Tyr (Gly Gly Ser)

RP100:

(Ser Gly Gly) Thr Arg Lys Gly Ile His Ile Gly Pro Gly Arg Ala Ile Tyr (Gly Gly Ser Cys)

RP102:

(Ser Gly Gly) Thr Arg Lys Ser Ile Ser Ile Gly Pro Gly Arg Ala Phe (Gly Gly Ser Cys)

RP91 (MN-Hepatitis fusion):

Arg Ile His Ile Gly Pro Gly Arg Ala Phe Tyr Gly Phe Phe Leu Leu Thr Arg Ile Leu Thr Ile Pro Gln Ser Leu Asp (Cys)

RP104:

(Ser Gly Gly) Ile Gly Pro Gly Arg Ala Phe Tyr Thr Thr Lys (Gly Gly Ser Cys)

RP106:

(Ser Gly Gly) Arg Ile His Ile Gly Pro Gly Arg Ala Phe (Gly Gly Ser Cys)

RP108:

(Ser Gly Gly) His Ile Gly Pro Gly Arg Ala Phe Tyr Ala Thr Gly (Gly Gly Ser Cys)

RP70 (from isolate HIV-MN):

Ile Asn Cys Thr Arg Pro Asn Tyr Asn Lys Arg Lys Arg Ile His Ile Gly Pro Gly Arg Ala Phe Tyr Thr Thr Lys Asn Ile Ile Gly Thr Ile Arg Gln Ala His Cys Asn Ile Ser

RP84 (from isolate HIV-MN):

Ile His Ile Gly Pro Gly Arg Ala Phe Tyr Thr Gly (Cys)

RP144 (from isolate 7887-3):

Asn Asn Thr Ser Arg Gly Ile Arg Ile Gly Pro Gly Arg Ala Ile Leu Ala Thr Glu Arg Ile Ile Gly (Cys)

RP145 (from isolate 6587-7):

Asn Asn Thr Arg Lys Gly Ile His Ile Gly Pro Gly Arg Ala Phe Tyr Ala Thr Gly Asp Ile Ile Gly (Cys)

RP146 (from isolate CC):

Asn Asn Thr Lys Lys Gly Ile Arg Ile Gly Pro Gly Arg Ala
Val Tyr Thr Ala Arg Arg Ile Ile Gly (Cys)

RP147 (from isolate KK261):

Asn Asn Thr Arg Lys Gly Ile Tyr Val Gly Ser Gly Arg Lys
Val Tyr Thr Arg His Lys Ile Ile Gly (Cys)

RP150 (from isolate ARV-2):

Asn Asn Thr Arg Lys Ser Ile Tyr Ile Gly Pro Gly Arg Ala
Phe His Thr Thr Gly Arg Ile Ile Gly (Cys)

RP151 (from isolate NY5):

Asn Asn Thr Lys Lys Gly Ile Ala Ile Gly Pro Gly Arg Thr
Leu Tyr Ala Arg Glu Lys Ile Ile Gly (Cys)

C. Hybrid Peptides Containing Sequences from More Than One Variant

RP73 (from isolates HIV-IIIB, HIV-RF):

Lys Ser Ile Arg Ile Gln Arg Gly Pro Gly Arg Val Ile
Tyr (Cys)

RP74 (from isolates HIV-IIIB, HIV-RF, HIV-MN, HIV-SC):

Arg Ile His Ile Gly Pro Gly Arg Ala Phe Tyr Ala Lys
Ser Ile Arg Ile Gln Arg Gly Pro Gly Arg Val Ile Tyr
(Cys)

RP80 (from isolates HIV-IIIB, HIV-RF):

Arg Ile Gln Arg Gly Pro Gly Arg Val Ile Tyr Ala Thr
(Cys)

RP81 (from isolates HIV-IIIB, HIV-RF, HIV-WMJ1, HIV-MN):

Arg Ile His Ile Gly Pro Gly Arg Ala Phe Tyr Thr Gly
Arg Ile Gln Arg Gly Pro Gly Arg Val Ile Tyr Ala Thr
(Cys)

RP82 (from isolates HIV-MN, HIV-WMJ1):

Arg Ile His Ile Gly Pro Gly Arg Ala Phe Tyr Thr Gly
(Cys)

RP88 (from isolates HIV-MN, HIV-SC):

Ser Ile His Ile Gly Pro Gly Arg Ala Phe Tyr Thr Thr Gly
(Cys)

RP137 (from isolates HIV-IIIB, HIV-RF):

Asn Asn Thr Arg Lys Ser Ile Arg Ile Thr Lys Gly Pro
Gly Arg Ala Phe Val Thr Ile Gly Lys Ile Gly (Cys)

RP140 (from isolates HIV-IIIB, HIV-RF):

Asn Asn Thr Arg Lys Ser Ile Thr Lys Gly Pro Gly Arg
Ala Phe Val Thr Ile Gly Lys Ile Gly (Cys)

Peptide 64 (from isolates HIV-IIIB, HIV-RF, HIV-MN, HIV-SC):

Arg Ile His Ile Gly Pro Gly Arg Ala Ile Phe Tyr Arg
Ile Gln Arg Gly Pro Gly Arg Val Ile Tyr (Cys)

Peptide 338 (from isolates HIV-IIIB, HIV-RF):

Arg Ile Gln Arg Gly Pro Gly Arg Val Ile Tyr (Cys)

Peptide 138 (from isolates HIV-IIIB, HIV-RF):

Asn Asn Thr Arg Lys Ser Ile Arg Ile Gln Arg Gly
Pro Gly Arg Val Ile Tyr Ala Thr Gly Lys Ile Gly
(Cys)

RP63 (III_B-RF hybrid):

Arg Ile Gln Arg Gly Pro Gly Arg Val Ile Tyr (Cys)

D. Miscellaneous Peptide Sequences

RP41:

Gly Pro Gly Arg

RP61:

Gly Pro Gly Arg (Ala Ala Ala Ala Ala Ala Cys)

RP75:

(Cys Ala Ala Ala Ala Ala) Gly Pro Gly Arg Ala Phe (Ala Ala Ala Cys)

RP111:

Ile Gln Arg Gly Pro Gly Ile Gln Arg Gly Pro Gly (Cys)

RP113:

Gln Arg Gly Pro Gly Arg Gln Arg Gly Pro Gly Arg Gln Arg Gly Pro Gly Arg (Cys)

RP114:

Arg Gly Pro Gly Arg Gly Pro Gly Arg Gly Pro Gly Arg Gly Pro Gly Arg (Cys)

RP116:

Gly Pro Gly Arg Ala Phe Gly Pro Gly Arg Ala Phe Gly Pro Gly Arg Ala Phe (Cys)

RP120:

Ser Ile Arg Ile Gly Pro Gly Arg Ala Phe Tyr Thr (Cys)

RP121c:

(Cys) Gly Pro Gly Arg (Cys)

RP122c:

(Cys) Ile Gly Pro Gly Arg Ala (Cys)

RP123c:

(Cys) His Ile Gly Pro Gly Arg Ala Phe (Cys)

The proteins and peptides exemplifying the subject invention can be made by well-known synthesis procedures. Alternatively, these entities can be made by use of recombinant DNA procedures. Such recombinant DNA procedures are disclosed herein since they were, in fact, the procedures initially utilized to obtain the novel proteins and peptides of the invention. However, once these entities were prepared and their molecules sequenced, it is apparent to a person skilled in the art that the preferred method for making them would now be by chemical synthesis means. For example, there are available automated machines which can readily make proteins and peptides of the molecular sizes disclosed herein.

In the recombinant DNA procedures for making some of the proteins and peptides of the invention, an expression vector plasmid denoted pREV2.2 was used. This plasmid was initially constructed from a plasmid denoted pBG1.

Plasmid pBG1 is deposited in the E. coli host MS371 with the Northern Regional Research Laboratory (NRRL, U.S. Department of Agriculture, Peoria, Illinois, USA). It is in the permanent collection of this repository. E. coli MS371(pBG1), NRRl B-15904, was deposited on November 1, 1984. E. coli MS371, NRRL B-15129 is now available to the public.

Plasmid pREV2.2 was deposited in the E. coli JM103 host with NRRL on July 30, 1986. E. coli JM103(pREV2.2) received the accession number NRRL B-18091. NRRL B-15904 and NRRL B-18091 will be available, without restrictions, to the public upon the grant of a patent which discloses them.

Other E. coli strains, disclosed herein, were deposited as follows:

E. coli SG20251, NRRL B-15918, was deposited on December 12, 1984.

E. coli CAG629(pKH1), NRRL B-18095, was deposited on July 30, 1986.

This latter deposit can be subjected to standard techniques to separate the plasmid from the host cell, and, thus, use the host E. coli CAG629 as disclosed herein.

The subject cultures have been deposited under conditions that assure that access to the cultures will be available during the pendency of the patent application disclosing them to one determined by the Commissioner of Patents and Trademarks to be entitled thereto under 37 CFR 1.14 and 35 USC 122. The deposits are available as required by foreign patent laws in countries wherein counterparts of the subject application, or its progeny, are filed. However, it should be understood that the availability of a deposit does not constitute a license to practice the subject invention in derogation of patent rights granted by governmental action.

Further, the subject culture deposits will be stored and made available to the public in accord with the provisions of the Budapest Treaty for the Deposit of Microorganisms, i.e., they will be stored with all the care necessary to keep them viable and uncontaminated for a period of at least five years after the most recent request for the furnishing of a sample of the deposits, and in any case, for a

period of at least 30 (thirty) years after the date of deposit or for the enforceable life of any patent which may issue disclosing the cultures. The depositor acknowledges the duty to replace the deposits should the depository be unable to furnish a sample when requested, due to the condition of the deposits. All restrictions on the availability to the public of the subject culture deposits will be irrevocably removed upon the granting of a patent disclosing them.

The novel HIV proteins and peptides of the subject invention can be expressed in Saccharomyces cerevisiae using plasmids containing the inducible galactose promoter from this organism (Broach, J.R., Y. Li, L.C. Wu, and M. Jayaram [1983] in Experimental Manipulation of Gene Expression, p. 83, ed. M. Inouye, Academic Press). These plasmids are called YEp51 and YEp52 (Broach, J.R. et al [1983]) and contain the E. coli origin of replication, the gene for β-lactamase, the yeast LEU2 gene, the 2 μm origin of replication and the 2 μm circle REP3 locus. Recombinant gene expression is driven by the yeast GAL10 gene promoter.

Yeast promoters such as galactose and alcohol dehydrogenase (Bennetzen, J.L. and B.D. Hall [1982] J. Biol. Chem. 257:3018; Ammerer, G. [1983] in Methods in Enzymology Vol. 101, p. 192), phosphoglycerate kinase (Derynck, R., R.A. Hitzeman, P.W. Gray, D.V. Goeddel [1983] in Experimental Manipulation of Gene Expression, p. 247, ed. M. Inouye, Academic Press), triose phosphate isomerase (Alber, T. and G. Kawasaki [1982] J. Molec. and Applied Genet. 1:419), or enolase (Innes, M.A. et al. [1985] Science 226:21) can be used.

The genes disclosed herein can be expressed in simian cells. When the genes encoding these proteins are cloned into one of the plasmids as described in Okayama and Berg (Okayama, H. and P. Berg [1983] Molec. and Cell. Biol. 3:280) and references therein, or COS cells transformed with these plasmids, synthesis of HIV proteins can be detected immunologically.

Other mammalian cell gene expression/protein production systems can be used. Examples of other such systems are the vaccinia virus expression system (Moss, B. [1985] Immunology Today 6:243; Chakrabarti, S., K. Brechling, B. Moss [1985] Molec. and Cell. Biol. 5:3403) and the vectors derived from murine retroviruses (Mulligan, R.C. [1983] in Experimental Manipulation of Gene Expression, p. 155, ed. M. Inouye, Academic Press).

The HIV proteins and peptides of the subject invention can be chemically synthesized by solid phase peptide synthetic techniques such as BOC and FMOC (Merrifield, R.B. [1963] J. Amer. Chem. Soc. 85:2149; Chang, C. and J. Meienhofer [1978] Int. J. Peptide Protein Res. 11:246).

As is well known in the art, the amino acid sequence of a protein is determined by the nucleotide sequence of the DNA. Because of the redundancy of the genetic code, i.e., more than one coding nucleotide triplet (codon) can be used for most of the amino acids used to make proteins, different nucleotide sequences can code for a particular amino acid. Thus, the genetic code can be depicted as follows:

Phenylalanine (Phe)	TTK	Histidine (His)	CAK
Leucine (Leu)	XTY	Glutamine (Gln)	CAJ
Isoleucine (Ile)	ATM	Asparagine (Asn)	AAK
Methionine (Met)	ATG	Lysine (Lys)	AAJ
Valine (Val)	GTL	Aspartic acid (Asp)	GAK
Serine (Ser)	QRS	Glutamic acid (Glu)	GAJ
Proline (Pro)	CCL	Cysteine (Cys)	TGK
Threonine (Thr)	ACL	Tryptophan (Trp)	TGG
Alanine (Ala)	GCL	Arginine (Arg)	WGZ
Tyrosine (Tyr)	TAK	Glycine (Gly)	GGL
Termination signal	TAJ		
Termination signal	TGA		

Key: Each 3-letter deoxynucleotide triplet corresponds to a trinucleotide of mRNA, having a 5'-end on the left and a 3'-end on the right. All DNA sequences given herein are those of the strand whose sequence corresponds to the mRNA sequence, with thymine substituted for uracil. The letters stand for the purine or pyrimidine bases forming the deoxynucleotide sequence.

A = adenine
G = guanine
C = cytosine
T = thymine
X = T or C if Y is A or G
X = C if Y is C or T
Y = A, G, C or T if X is C
Y = A or G if X is T
W = C or A if Z is A or G
W = C if Z is C or T
Z = A, G, C or T if W is C
Z = A or G if W is A
QR = TC if S is A, G, C or T; alternatively
QR = AG if S is T or C
J = A or G
K = T or C
L = A, T, C or G
M = A, C or T

The above shows that the novel amino acid sequences of the HIV proteins and peptides of the subject invention can be prepared by nucleotide sequences other than those disclosed herein. Functionally equivalent nucleotide sequences encoding the novel amino acid sequences of these HIV proteins and peptides, or fragments thereof having HIV antigenic or immunogenic or therapeutic activity, can be prepared by known synthetic procedures. Accordingly, the subject invention includes such functionally equivalent nucleotide sequences.

Thus the scope of the subject invention includes not only the specific nucleotide sequences depicted herein, but also all equivalent nucleotide sequences coding for molecules with substantially the same HIV antigenic or immunogenic or therapeutic activity.

Further, the scope of the subject invention is intended to cover not only the specific amino acid sequences disclosed, but also similar sequences coding for proteins or protein fragments having comparable ability to induce the formation of and/or bind to specific HIV antibodies possessing the properties of virus neutralization.

The term "equivalent" is being used in its ordinary patent usage here as denoting a nucleotide sequence which performs substantially as the nucleotide sequence identified herein to produce molecules with substantially the same HIV antigenic or immunogenic or therapeutic activity in essentially the same kind of hosts. Within this definition are subfragments which have HIV antigenic or immunogenic or therapeutic activity.

As disclosed above, it is well within the skill of those in the genetic engineering art to use the nucleotide sequences encoding HIV antigenic or immunogenic or therapeutic activity of the subject invention to produce HIV proteins via microbial processes. Fusing the sequences into an expression vector and transforming or transfecting into hosts, either eukaryotic (yeast or mammalian cells) or prokaryotic (bacterial cells), are standard procedures used in producing other well-known proteins, e.g., insulin, interferons, human growth hormone, IL-1, IL-2, and the like. Similar procedures, or obvious modifications thereof, can be employed to prepare HIV proteins or peptides by microbial means or tissue-culture technology in accord with the subject invention.

The nucleotide sequences disclosed herein can be prepared by a "gene machine" by procedures well known in the art. This is possible because of the disclosure of the nucleotide sequence.

The restriction enzymes disclosed can be purchased from Bethesda Research Laboratories, Gaithersburg, MD, or New England Biolabs, Beverly, MA. The enzymes are used according to the instructions provided by the supplier.

The various methods employed in the preparation of the plasmids and transformation of host organisms are well known in the art. These procedures are all described in Maniatis, T., E.F. Fritsch, and J. Sambrook (1982) Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory, New York. Thus, it is within the skill of those in the genetic engineering art to extract DNA from

microbial cells, perform restriction enzyme digestions, electrophorese DNA fragments, tail and anneal plasmid and insert DNA, ligate DNA, transform cells, e.g., E. coli cells, prepare plasmid DNA, electrophorese proteins, and sequence DNA.

Immunochemical assays employing the HIV proteins or peptides of the invention can take a variety of forms. One preferred type is a liquid phase assay wherein the HIV antigen and the sample to be tested are mixed and allowed to form immune complexes in solution which are then detected by a variety of methods. Another preferred type is a solid phase immunometric assay. In solid phase assays, an HIV protein or peptide is immobilized on a solid phase to form an antigen-immunoadsorbent. The immunoadsorbent is incubated with the sample to be tested. After an appropriate incubation period, the immunoadsorbent is separated from the sample, and labeled anti-(human IgG) antibody is used to detect human anti-HIV antibody bound to the immunoadsorbent. The amount of label associated with the immunoadsorbent can be compared to positive and negative controls to assess the presence or absence of anti-HIV antibody.

The immunoadsorbent can be prepared by adsorbing or coupling a purified HIV protein or peptide to a solid phase. Various solid phases can be used, such as beads formed of glass, polystyrene, polypropylene, dextran or other material. Other suitable solid phases include tubes or plates formed from or coated with these materials.

The HIV proteins or peptides can be either covalently or non-covalently bound to the solid phase by techniques such as covalent bonding via an amide or ester linkage or adsorption. After the HIV protein or peptide is affixed to the solid phase, the solid phase can be post-coated with an animal protein, e.g., 3% fish gelatin. This provides a blocking protein which reduces nonspecific adsorption of protein to the immunoadsorbent surface.

The immunoadsorbent is then incubated with the sample to be tested for anti-HIV antibody. In blood screening, blood plasma or serum is used. The plasma or serum is diluted with normal animal plasma or serum. The diluent plasma or serum is derived from the same animal species that is the source of the anti-(human IgG) antibody. The preferred anti-(human IgG) antibody is goat anti-(human IgG) antibody. Thus, in the preferred format, the diluent would be goat serum or plasma.

The conditions of incubation, e.g., pH and temperature, and the duration of incubation are not crucial. These parameters can be optimized by routine experimentation. Generally, the incubation will be run for 1-2 hr at about 45°C in a buffer of pH 7-8.

After incubation, the immunoadsorbent and the sample are separated. Separation can be accomplished by any conventional separation technique such as sedimentation or centrifugation. The immunoadsorbent then may be washed free of sample to eliminate any interfering substance.

The immunoadsorbent is incubated with the labeled anti-(human IgG) antibody (tracer) to detect human antibody bound thereto. Generally the immunoadsorbent is incubated with a solution

of the labeled anti-(human IgG) antibody which contains a small amount (about 1%) of the serum or plasma of the animal species which serves as the source of the anti-(human IgG) antibody. Anti-(human IgG) antibody can be obtained from any animal source. However, goat anti-(human IgG) antibody is preferred. The anti-(human IgG) antibody can be an antibody against the Fc fragment of human IgG, for example, goat anti-(human IgG) Fc antibody.

The anti-(human IgG) antibody or anti-(human IgG) Fc can be labeled with a radioactive material such as ^{125}I; labeled with an optical label, such as a fluorescent material; or labeled with an enzyme such as horseradish peroxidase. The anti-human antibody can also be biotinylated and labeled avidin used to detect its binding to the immunoadsorbent.

After incubation with the labeled antibody, the immunoadsorbent is separated from the solution and the label associated with the immunoadsorbent is evaluated. Depending upon the choice of label, the evaluation can be done in a variety of ways. The label may be detected by a gamma counter if the label is a radioactive gamma emitter, or by a fluorimeter, if the label is a fluorescent material. In the case of an enzyme, label detection may be done colorimetrically employing a substrate for the enzyme.

The amount of label associated with the immunoadsorbent is compared with positive and negative controls in order to determine the presence of anti-HIV antibody. The controls are generally run concomitantly with the sample to be tested. A positive control is a serum containing antibody against HIV; a negative control is a serum from healthy individuals which does not contain antibody against HIV.

For convenience and standardization, reagents for the performance of the immunometric assay can be assembled in assay kits. A kit for screening blood, for example, can include:

(a) an immunoadsorbent, e.g., a polystyrene bead coated with an HIV protein or peptide;

(b) a diluent for the serum or plasma sample, e.g. normal goat serum or plasma;

(c) an anti-(human IgG) antibody, e.g., goat anti-(human IgG) antibody in buffered, aqueous solution containing about 1% goat serum or plasma;

(d) a positive control, e.g., serum containing antibody against at least one of the novel HIV proteins or peptides; and

(e) a negative control, e.g., pooled sera from healthy individuals which does not contain antibody against at least one of the novel HIV proteins or peptides.

If the label is an enzyme, an additional element of the kit can be the substrate for the enzyme.

Another type of assay for anti-HIV antibody is an antigen sandwich assay. In this assay, a labeled HIV protein or peptide is used in place of anti-(human IgG) antibody to detect anti-HIV antibody bound to the immunoadsorbent. The assay is based in principle on the bivalency of antibody molecules. One binding site of the antibody binds the antigen affixed to the solid phase; the second

is available for binding the labeled antigen. The assay procedure is essentially the same as described for the immunometric assay except that after incubation with the sample, the immunoadsorbent is incubated with a solution of labeled HIV protein or peptide. HIV proteins or peptides can be labeled with radioisotope, an enzyme, etc. for this type of assay.

In a third format, the bacterial protein, protein A, which binds the Fc segment of an IgG molecule without interfering with the antigen-antibody interaction can be used as the labeled tracer to detect anti-HIV antibody adsorbed to the immunoadsorbent. Protein A can be readily labeled with a radioisotope, enzyme, or other detectable species.

Immunochemical assays employing an HIV protein or peptide have several advantages over those employing a whole (or disrupted) virus. Assays based upon an HIV protein or peptide will alleviate the concern over growing large quantities of infectious virus and the inherent variability associated with cell culturing and virus production. Further, the assay will help mitigate the real or perceived fear of contracting AIDS by technicians in hospitals, clinics and blood banks who perform the test.

Immunochemical assays employing recombinant envelope proteins from multiple viral variants have additional advantages over proteins from a single HIV variant. Assays incorporating protein sequences from multiple variants are more likely to accurately survey antibodies in the human population infected with diverse HIV variants. Also, solid phase enzyme-linked immunosorbent assay (ELISA) utilizing different HIV variant proteins would allow determination of prevalent serotypes in different geographic locations. This determination has not been possible until now as no available antibody detection kit utilizes more than one HIV variant.

Another use of recombinant proteins from HIV variants is to elicit variant-specific antisera in test animals. This antiserum would provide a reagent to identify which viral variant infected an individual. Currently, "virus typing" can only be done by viral gene cloning and sequencing. Binding of variant-specific serum to a patient viral isolate would provide a means of rapid detection not currently available. For example, sera raised to the proteins denoted $PB1_{IIIB}$, $PB1_{RF}$, $PB1_{MN}$, $PB1_{SC}$, and $PB1_{WMJ2}$ can be used to screen viral isolates from patients to determine which HIV variant the clinical isolate most closely resembles. This "screening" can be done by a variety of known antibody-antigen binding techniques.

Vaccines comprising one or more of the HIV proteins or peptides, disclosed herein, and variants thereof having antigenic properties, can be prepared by procedures well known in the art. For example, such vaccines can be prepared as injectables, e.g., liquid solutions or suspensions. Solid forms for solution in, or suspension in, a liquid prior to injection also can be prepared. Optionally, the preparation also can be emulsified. The active antigenic ingredient or ingredients can be mixed with excipients which are pharmaceutically acceptable and compatible with the active ingredient.

Examples of suitable excipients are water, saline, dextrose, glycerol, ethanol, or the like, and combinations thereof. In addition, if desired, the vaccine can contain minor amounts of auxiliary substances such as wetting or emulsifying agents, pH buffering agents, or adjuvants such as aluminum hydroxide or muramyl dipeptide or variations thereof. In the case of peptides, coupling to larger molecules such as KLH sometimes enhances immunogenicity. The vaccines are conventionally administered parenterally, by injection, for example, either subcutaneously or intramuscularly. Additional formulations which are suitable for other modes of administration include suppositories and, in some cases, oral formulations. For suppositories, traditional binders and carriers include, for example, polyalkalene glycols or triglycerides. Suppositories can be formed from mixtures containing the active ingredient in the range of about 0.5% to about 10%, preferably about 1 to about 2%. Oral formulations can include such normally employed excipients as, for example, pharmaceutical grades of mannitol, lactose, starch, magnesium stearate, sodium saccharine, cellulose, magnesium carbonate, and the like. These compositions can take the form of solutions, suspensions, tablets, pills, capsules, sustained release formulations or powders and contain from about 10% to about 95% of active ingredient, preferably from about 25% to about 70%.

The compounds can be formulated into the vaccine as neutral or salt forms. Pharmaceutically acceptable salts include the acid addition salts (formed with the free amino groups of the peptide) and which are formed with inorganic acids such as, for example, hydrochloric or phosphoric acids, or such organic acids as acetic, oxalic, tartaric, mandelic, and the like. Salts formed with the free carboxyl groups can also be derived from inorganic bases such as, for example, sodium, potassium, ammonium, calcium, or ferric hydroxides, and such organic bases as isopropylamine, trimethylamine, 2-ethylamino ethanol, histidine, procaine, and the like. A vaccine composition may include peptides containing T helper cell epitopes in combination with protein fragments containing the principal neutralizing domain. Several of these epitopes have been mapped within the HIV envelope, and these regions have been shown to stimulate proliferation and lymphokine release from lymphocytes. Providing both of these epitopes in a vaccine may result in the stimulation of both the humoral and the cellular immune responses.

Alternatively, a vaccine composition may include a compound which functions to increase the general immune response. One such compound is interleukin-2 (IL-2) which has been reported to enhance immunogenicity by general immune stimulation (Nunberg et al. [1988] In New Chemical and Genetic Approaches to Vaccination, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY). IL-2 may be coupled with an HIV peptide or protein comprising the PND to enhance the efficacy of vaccination.

The vaccines are administered in a manner compatible with the dosage formulation, and in such amount as will be therapeutically effective and immunogenic. The quantity to be administered

depends on the subject to be treated, capacity of the subject's immune system to synthesize antibodies, and the degree of protection desired. Precise amounts of active ingredient required to be administered depend on the judgment of the practitioner and are peculiar to each individual. However, suitable dosage ranges are of the order of about several hundred micrograms active ingredient per individual. Suitable regimes for initial administration and booster shots are also variable, but are typified by an initial administration followed in one or two week intervals by a subsequent injection or other administration.

HIV is known to undergo amino acid sequence variation, particularly in the envelope gene (Starcich, B.R. [1986] Cell 45:637-648; Hahn, B.H. et al. [1986] Science 232:1548-1553). Over 100 variants have been analyzed by molecular cloning and restriction enzyme recognition analysis, and several of these have been analyzed by nucleotide sequencing. Some of these are the HIV isolates known as RF (Popovic, M. et al. [1984] Science 224:497-500), WMJ-1 (Hahn, B.H. et al. [1986] Science 232:1548-1553), LAV (Wain-Hobson, S. et al. [1985] Cell 40:9-17), and ARV-2 (Sanchez-Pescador, R. et al. [1985] Science 227:484-492). HIV peptides from different viral isolates can be used in vaccine preparations to protect against infections by different HIV isolates. Further, a vaccine preparation can be made using more than one envelope protein fragment corresponding to the principal neutralizing domain of more than one HIV isolate to provide immunity and thus give better protection against AIDS. Alternatively, the vaccine preparation can be made using a single protein fragment that is comprised of a tandem arrangement of principal neutralizing epitopes from more than one HIV isolate. By identifying the principal neutralizing domain of HIV, this polypeptide region can be applied to formulate valuable vaccine, diagnostic, and therapeutic reagents.

Antibodies to recombinant peptides disclosed herein are useful as therapeutic and prophylactic reagents. The generation of polyclonal or monoclonal antibodies capable of neutralizing a variety of HIV variants could be used to reduce the incidence of accidental infection and treat HIV infected people that are immuno-compromised. Additionally, immunization regimens may elicit polyclonal sera capable of broadly neutralizing several variants of HIV. The ability to neutralize multiple HIV variants is termed broadly neutralizing antibody. Broadly neutralizing antibody may neutralize two or more HIV variants or all HIV variants. Therefore, a mixture of broadly neutralizing antibodies that neutralize different groups of HIV variants would be useful for diagnosis, prophylaxis, and therapy of AIDS.

It is surprising and advantageous that immunization with peptides from five HIV variants would yield sera capable of neutralizing more than these five HIV variants when immunization with two does not. Example 22 shows that immunization with five peptides elicits broadly neutralizing sera. Broadly neutralizing sera may also be generated if several sequences from the hypervariable region of diverse HIV variants are presented as a single synthetic peptide. Additionally, one may elicit this

broadly neutralizing sera by reimmunization of animals primed with RP136 or equivalent proteins with peptides containing only the conserved amino acids within this hypervariable region. These immunization regimens would be useful for vaccination and for deriving antibodies useful as therapeutic agents.

Polyvalent immune globulin for use in passive immunization can be prepared by immunization of horses or by pooling immune human sera and fractionation of the IgG component from plasma or sera. Human or mouse monoclonal antibody producing cell lines may be prepared by standard transformation and hybridoma technology (Methods in Enzymology, Vol. 121, Sections I and II [1986] eds. J.J. Langone and H.V. Vunakis, Academic Press). HIV monoclonal antibody can be prepared in accord with the procedures disclosed by Matsushita et al. (Matsushita et al. [1988] Journal of Virology 62(6):2107-2114). Since, for the most part, monoclonal antibodies are produced in species other than humans, they are often immunogenic to humans. In order to successfully use these monoclonal antibodies in the treatment of humans, it may be necessary to create a chimeric antibody molecule wherein the portion of the polypeptide involved with ligand binding (the variable region) is derived from one species, and the portion involved with providing structural stability and other biological functions (the constant region) is derived from a human antibody. Methods for producing chimeric antibodies in which the variable domain is derived from one host and the constant domain is derived from a second host are well known to those skilled in the art. See, for example, Neuberger et al., WO Publication No. 86/01533, priority 9/3/84; Morrison et al., EP Publication No. 0 173 494, priority 8/27/84. An alternative method, in which an antibody is produced by replacing only the complementarity determining regions (CDRs) of the variable region with the CDRs from an immunoglobulin of the desired antigenic specificity, is described by Winter (GB Publication No. 2 188 638, priority 3/27/86). Murine monoclonals can be made compatible with human therapeutic use by producing an antibody containing a human Fc portion (Morrison, S.L. [1985] Science 229:1202-1207). Established procedures would allow construction, expression, and purification of such a hybrid monoclonal antibody. Regimens for administering immune globulin therapeutically have previously been used for a number of infectious diseases.

As used herein, the term "antibody" is meant to encompass monoclonal or polyclonal antibodies, whole, intact antibodies or antibody fragments having the immunological activity of the whole antibody. Also encompassed within the term "antibody" are chimeric antibodies having the variable and constant regions from different host species, or those wherein only the CDRs are replaced.

For treatment of HIV infection, compositions comprising antibodies may be administered to an individual or animal in need of treatment. Alternatively, the HIV antigens described here may be administered in order to stimulate the recipient's own immune response. When treating with an HIV antigen, a single antigen may be administered or, preferably, a broadly neutralizing antigen or mixture

of antigens may be administered. Such compositions are described in detail in the examples which follow.

The ability to modify peptides made by organic synthesis can be advantageous for diagnostic, therapeutic, and prophylactic use by improving efficiency of immobilization, increasing protein stability, increasing immunogenicity, altering immunogenicity, reducing toxicity, or allowing multiple variations simultaneously. For example, peptides can be modified to increase or decrease net charge by modification of amino or carboxyl groups (carbamylation, trifluoroacetylation or succinylation of amino groups; acetylation of carboxyl groups). Peptides can be made more stable by, for example, inclusion of D-amino acids or circularization of the peptide. Reductive state of peptides can be altered by, for example, sulfonation of cystinyl groups. Peptides can also be modified covalently or non-covalently with non-proteinaceous materials such as lipids or carbohydrates to enhance immunogenicity or solubility. Polyethylene glycol can be used to enhance solubility. The subject invention includes all such chemical modifications of the proteins and peptides disclosed herein so long as the modified protein and/or peptide retains substantially all the antigenic/immunogenic properties of the parent compound.

Peptides can also be modified to contain antigenic properties of more than one viral variant. This has been done, for example, with Foot and Mouth Disease virus.

Foot and Mouth Disease virus is similar to HIV in that multiple variants exist and immunization with one variant does not lead to protection against other variants. The real utility of peptides as immunogens is demonstrated by eliciting immunity to more than one variant by modification of the peptide to possess properties of both natural variants. When such a modified variant was used to immunize test animals, they were protected against both Foot and Mouth virus strains A10 and A12 (Brown, F. in Virus Vaccines, ed. G. Dreesman, J. Bronson, R. Kennedy, pp. 49-54 [1985]).

HIV peptides or proteins containing a PND epitope can also be coupled with or incorporated into an unrelated virus particle, a replicating virus, or other microorganism in order to enhance immunogenicity. The HIV epitope may be genetically or chemically attached to the viral particle or microorganism or an immunogenic portion or component thereof. Antigenic epitopes have been attached to viral proteins or particles to enhance the immune response. For example, the VP6 capsid protein of rotavirus has been used as an immunologic carrier protein for an epitope of interest either in the monomeric form or as oligomers of VP6 in the form of particles (EP Publication No. 0 259 149). Similarly, Evans et al. (1989, Nature 339:385) have constructed chimaeras of the poliovirus capsid protein and an epitope of HIV gp41 to enhance immunogenicity of the HIV epitope. Foreign antigenic determinants have also been expressed and presented by bacterial cells. A Salmonella strain expressing a cloned Salmonella flagellin gene, into which was inserted an epitope of either cholera

toxin or hepatitis B surface antigen, was reported to elicit both cellular and humoral responses to the inserted epitopes (Newton et al. [1989] Science 244:70-72; and Wu et al. [1989] Proc. Natl. Acad. Sci. 86:4726-4730).

Example 18 shows that a peptide containing amino acid sequences from two HIV variants can block virus neutralization activity of two virus specific neutralizing antisera. This suggests that a peptide or protein containing sequences of two or more HIV variants can elicit an immune response effective against two or more HIV variants.

Example 19 shows that co-immunization with envelope proteins from two HIV isolates elicits an immune response capable of neutralizing two HIV isolates. This suggests that co-immunization with proteins from two or more HIV variants can elicit an immune response effective against two or more HIV variants.

Following are examples which illustrate the process of the invention, including the best mode. These examples should not be construed as limiting. All solvent mixture proportions are by volume unless otherwise noted.

Example 1 – Construction of plasmid pREV2.2

The pREV2.2 plasmid expression vector was constructed from plasmid pBG1. Plasmid pBG1 can be isolated from its E. coli host by well known procedures, e.g., using cleared lysate-isopycnic density gradient procedures, and the like. Like pBG1, pREV2.2 expresses inserted genes behind the E. coli promoter. The differences between pBG1 and pREV2.2 are the following:

1. pREV2.2 lacks a functional replication of plasmid (rop) protein.
2. pREV2.2 has the trpA transcription terminator inserted into the AatII site. This sequence insures transcription termination of over-expressed genes.
3. pREV2.2 has genes to provide resistance to ampicillin and chloramphenicol, whereas pBG1 provides resistance only to ampicillin.
4. pREV2.2 contains a sequence encoding sites for several restriction endonucleases.

The following procedures were used to make each of the four changes listed above:

1a. 5 μg of plasmid pBG1 was restricted with NdeI, which gives two fragments of approximately 2160 and 3440 base pairs.

1b. 0.1 μg of DNA from the digestion mixture, after inactivation of the NdeI, was treated with T4 DNA ligase under conditions that favor intramolecular ligation (200 μl reaction volume using standard T4 ligase reaction conditions [New England Biolabs, Beverly, MA]). Intramolecular ligation of the 3440 base pair fragment gave an ampicillin resistant plasmid. The ligation mixture was transformed into the recipient

strain E. coli JM103 (available from New England Biolabs) and ampicillin resistant clones were selected by standard procedures.

1c. The product plasmid, pBG1ΔN, where the 2160 base pair NdeI fragment is deleted from pBG1, was selected by preparing plasmid from ampicillin resistant clones and determining the restriction digestion patterns with NdeI and SalI (product fragments approximately 1790 and 1650). This deletion inactivates the rop gene that controls plasmid replication.

2a. 5 μg of pBG1 N was then digested with EcoRI and BclI and the larger fragment, approximately 2455 base pairs, was isolated.

2b. A synthetic double stranded fragment was prepared by the procedure of Itakura et al. (Itakura, K., J.J. Rossi, and R.B. Wallace [1984] Ann. Rev. Biochem. 53:323-356, and references therein) with the following structure:

```
5'    GATCAAGCTTCTGCAGTCGACGCAT
3'    TTCGAAGACGTCAGCTGCGTACGCC
                              GCGGATCCGGTACCCGGGAGCTCG 3'
                              TAGGCCATGGGCCCTCGAGCTTAA 5'
```

This fragment has BclI and EcoRI sticky ends and contains recognition sequences for several restriction endonucleases.

2c. 0.1 μg of the 2455 base pair EcoRI-BclI fragment and 0.01 μg of the synthetic fragment were joined with T4 DNA ligase and competent cells of strain JM103 were transformed. Cells harboring the recombinant plasmid, where the synthetic fragment was inserted into pBG1ΔN between the BclI and EcoRI sites, were selected by digestion of the plasmid with HpaI and EcoRI. The diagnostic fragment sizes are approximately 2355 and 200 base pairs. This plasmid is called pREV1.

2d. 5 μg of pREV1 were digested with AatII, which cleaves uniquely.

2e. The following double-stranded fragment was synthesized:

```
5'    CGGTACCAGCCCGCCTAATG
3'    TGCAGCCATGGTCGGGCGGA
                              AGCGGGCTTTTTTTTGACGT3'
                              TTACTCGCCCGAAAAAAAAC      5'
```

This fragment has AatII sticky ends and contains the trpA transcription termination sequence.

2f. 0.1 μg of AatII digested pREV1 was ligated with 0.01 μg of the synthetic fragment in a volume of 20 μl using T4 DNA ligase.

2g. Cells of strain JM103, made competent, were transformed and ampicillin resistant clones selected.

2h. Using a KpnI, EcoRI double restriction digest of plasmid isolated from selected colonies, a cell containing the correct construction was isolated. The sizes of the KpnI, EcoRI generated fragments are approximately 2475 and 80 base pairs. This plasmid is called pREV1TT and contains the trpA transcription terminator.

3a. 5 μg of pREV1TT, prepared as disclosed above (by standard methods) was cleaved with NdeI and XmnI and the approximately 850 base pair fragment was isolated.

3b. 5 μg of plasmid pBR325 (BRL, Gaithersburg, MD), which contains the genes conferring resistance to chloramphenicol as well as to ampicillin and tetracycline, was cleaved with BclI and the ends blunted with Klenow polymerase and deoxynucleotides. After inactivating the enzyme, the mixture was treated with NdeI and the approximately 3185 base pair fragment was isolated. This fragment contains the genes for chloramphenicol and ampicillin resistance and the origin of replication.

3c. 0.1 μg of the NdeI-XmnI fragment from pREV1TT and the NdeI-BclI fragment from pBR325 were ligated in 20 μl with T4 DNA ligase and the mixture used to transform competent cells of strain JM103. Cells resistant to both ampicillin and chloramphenicol were selected.

3d. Using an EcoRI and NdeI double digest of plasmid from selected clones, a plasmid was selected giving fragment sizes of approximately 2480, 1145, and 410 base pairs. This is called plasmid pREV1TT/chl and has genes for resistance to both ampicillin and chloramphenicol.

4a. The following double-stranded fragment was synthesized:

```
     MluI    EcoRV  ClaI BamHI  SalI HindIII SmaI
5'   CGAACGCGTGGCCGATATCATCGATGG-
                         ATCCGTCGACAAGCTTCCCGGGAGCT     3'
3'   GCTTGCGCACCGGCTATAGTAGCTAC-
                         CTAGGCAGCTGTTCGAAGGGCCC     5'
```

This fragment, with a blunt end and an SstI sticky end, contains recognition sequences for several restriction enzyme sites.

4b. 5 μg of pREV1TT/chl was cleaved with NruI (which cleaves about 20 nucleotides from the BclI site) and SstI (which cleaves within the multiple cloning site). The larger fragment, approximately 3990 base pairs, was isolated from an agarose gel.

4c. 0.1 μg of the NruI-SstI fragment from pREV1TT/chl and 0.01 μg of the synthetic fragment were treated with T4 DNA ligase in a volume of 20 μl.

4d. This mixture was transformed into strain JM103 and ampicillin resistant clones were selected.

4e. Plasmid was purified from several clones and screened by digestion with MluI or ClaI. Recombinant clones with the new multiple cloning site will give one fragment when digested with either of these enzymes, because each cleaves the plasmid once.

4f. The sequence of the multiple cloning site was verified. This was done by restricting the plasmid with HpaI and PvuII and isolating the 1395 base pair fragment, cloning it into the SmaI site of mp18 and sequencing it by dideoxynucleotide sequencing using standard methods.

4g. This plasmid is called pREV2.2.

Example 2 – Construction of the bacterial expression vector pREV2.1

Plasmid pREV2.1 was constructed using plasmid pREV2.2 and a synthetic oligonucleotide. The resulting plasmid was used to construct pPB1-Sub 1 and pPB1-Sub 2.

An example of how to construct pREV2.1 is as follows:

1. Plasmid pREV2.2 is cleaved with restriction enzymes NruI and BamHI and the 4 Kb fragment is isolated from an agarose gel.

2. The following double-stranded oligonucleotide is synthesized:

```
5'   CGAACGCGTGGTCCGATATCATCGATG      3'
3'   GCTTGCGCACCAGGCTATAGTAGCTACCTAG  5'
```

3. The fragments from 1 and 2 are ligated in 20 μl using T4 DNA ligase, transformed into competent E. coli cells and chloramphenicol resistant colonies are isolated.

4. Plasmid clones are identified that contain the oligonucleotide from 2. spanning the region from the NruI site to the BamHI site and recreating these two restriction sites. This plasmid is termed pREV2.1.

Example 3 – Construction of and expression from plasmid pPB1-Sub 1

Plasmid pPB1-Sub 1, which contains approximately 165 base pairs (bp) of DNA encoding essentially the HIV env gene from the PvuII site to the DraI site, and from which is synthesized an approximately 12 Kd fusion protein containing this portion of the gp120 envelope protein can be constructed as follows:

1. Restricting plasmid $pPB1_{IIIB}$ with MluI and DraI and isolating the approximately 165 bp fragment.

2. Restricting plasmid pREV2.1 with MluI and SmaI and isolating the large fragment, approximately 4 Kd, from an agarose gel.
3. Ligating the fragment prepared in 2. with the pREV2.1 fragment in a volume of 20 μl using T4 DNA ligase, transforming the ligation mixture into competent cell strain CAG629, and selecting ampicillin-resistant transformants.
4. Selecting such transformants, by appropriate restriction patterns, that have the gp120 fragment cloned in the proper orientation to generate a fusion protein. When the strain harboring this recombinant plasmid is grown in 2% medium containing 50 μg/ml ampicillin and the total complement of cellular proteins are electrophoresed on a SDS-polyacrylamide gel, a protein of approximately 12 Kd can be visualized by either coomassie blue staining or by Western blot analysis using as probe selected sera from HIV infected individuals.

Example 4 – Purification of recombinant protein containing HIV envelope sequences from plasmid pPB1-Sub 1

1. Growth of cells: Cells were grown in a 10-liter volume in a Chemap (Chemapec, Woodbury, NY) fermentor in 2% medium (2% yeast extract, bacto-tryptone, casamino acids [Difco, Detroit, MI], 0.2% potassium monobasic, 0.2% potassium dibasic, and 0.2% sodium dibasic). Fermentation temperature was 30°C, the pH was 6.8, and air was provided at 1 vvm. Plasmid selection was provided 20 μg/ml chloramphenicol. Typical cell yield (wet weight) was 30 g/l.
2. Cell lysis: 50 g, wet cell weight, of E. coli containing the recombinant HIV envelope fusion protein were resuspended in a final volume of 100 ml in 50 mM Tris-Cl pH 8.0, 5 mM potassium ethylenediaminetetraacetic acid (KEDTA), 5 mM dithiothreitol (DTT), 15 mM β-mercaptoethanol, 0.5% TRITON™X-100 (Pharmacia, Piscataway, NJ), and 5 mM phenylmethylsulfonyl fluoride (PMSF). The suspension was incubated for 30 min at room temperature.

   This material was lysed using a BEAD-BEATER™ (Biospec Products, Bartlesville, OK) containing an equal volume of 0.1-0.15 mm glass beads. The lysis was done for 6 min at room temperature in 1-min intervals. The liquid was separated from the beads and centrifuged for 2.5 hr at 20,000 xg. The supernatant was removed and the pellet was resuspended in 100 ml 8 M urea, 20 mM Tris-Cl pH 8.0, 5 mM DTT, 15 mM β-mercaptoethanol, 5 mM PMSF, and 1 mM KEDTA. The pellet was solubilized using a polytron homogenizer and centrifuged at 20,000 xg for 2 hr.

3. CM Chromatography: The dialysate was loaded onto a 100 ml column (2.5 cm x 20 cm) packed with CM FAST FLOW SEPHAROSE™ (Pharmacia) equilibrated in 8 M urea, 10 mM 4-(2-hydroxyethyl-1-piperazine ethane-sulfonic acid (HEPES) pH 6.5, 15 mM β-mercaptoethanol, and 1 mM KEDTA at room temperature. The column was washed with 200 ml equilibration buffer, and the protein eluted with a 1.0 liter linear gradient from 0-0.4 M NaCl. The HIV protein (12 Kd) eluted at approximately 0.2 M NaCl as assayed by SDS-polyacrylamide gel electrophoresis.

   Further purification was obtained by pooling Sub 1-containing fractions and applying to a S-200 (Pharmacia) gel filtration column equilibrated in the same buffer as the previous column.

Example 5 – Construction of and expression from plasmid pPB1-Sub 2

Plasmid pPB1-Sub 2, which contains approximately 320 bp of DNA encoding essentially the HIV env gene from the PvuII site to the ScaI site, and from which is synthesized an approximately 18 Kd fusion protein containing this portion of the gp120 envelope protein, can be constructed as follows:

1. Restricting the pPB1$_{IIIB}$ plasmid with Mlu1 and ScaI and isolating the approximately 320 bp fragment.
2. Restricting plasmid pREV2.1 with MluI and SmaI and isolating the large fragment, approximately 4 Kd, from an agarose gel.
3. Ligating the fragment prepared in 2. with the pREV2.1 fragment in a volume of 20 μl using T4 DNA ligase, transforming the ligation mixture into competent cell strain SG20251 and selecting ampicillin-resistant transformants.
4. Selecting such transformants, by appropriate restriction patterns, that have the gp120 fragment cloned in the proper orientation to generate a fusion protein. When the strain harboring this recombinant plasmid is grown in 2% medium containing 50 μg/ml ampicillin and the total complement of cellular proteins electrophoresed on an SDS-polyacrylamide gel, a protein of approximately 18 Kd can be visualized by either coomassie blue staining or by Western blot analysis using as probe selected sera from HIV infected individuals.

Example 6 – Purification of recombinant protein containing HIV envelope sequences from plasmid pPB1-Sub 2

1. Growth of cells: Cells were grown in a 10-liter volume in a Chemap fermentor in 2% medium. Fermentation temperature was 37°C, the pH was 6.8, and air was provided

at 1 vvm. Plasmid selection was provided by 20 μg/ml chloramphenicol. Typical cell yield (wet weight) was 30 g/l.

2. Cell lysis: 50 g, wet cell weight, of E. coli containing the recombinant HIV envelope fusion protein were resuspended in a final volume of 100 ml in 50 mM Tris-Cl pH 8.0, 5 mM KEDTA, 5 mM DTT, 15 mM β-mercaptoethanol, 0.5% TRITON™X-100, and 5 mM PMSF. The suspension incubated for 30 min at room temperature.

   This material was lysed using a BEAD-BEATER™ containing an equal volume of 0.1-0.15 mm glass beads. The lysis was done for 6 min at room temperature in 1 min intervals. The liquid was separated from the beads and centrifuged for 2.5 hr at 20,000 xg. The supernatant was removed and the pellet was resuspended in 100 ml 6 M guanidine-hydrochloride, 20 mM Tris-Cl pH 8.0, 5 mM DTT, 15 mM β-mercaptoethanol, 5 mM PMSF, and 1 mM KEDTA. The pellet was solubilized using a polytron homogenizer and centrifuged at 20,000 xg for 2 hr.

   The supernate (90 ml) was dialyzed against 4 liters of 8 M urea, 20 mM sodium formate, pH 4.0, 1 mM EDTA, and 15 mM β-mercaptoethanol. Dialysis was done each time for 8 hr or longer with three changes of buffer. Spectraphor dialysis tubing (S/P, McGraw Park, IL) with a 3.5 Kd MW cut-off was used.

3. CM Chromatography: The dialysate was loaded onto a 100 ml column (2.5 cm x 20 cm) packed with CM FAST FLOW SEPHAROSE™ (Pharmacia) equilibrated in 8 M urea, 20 mM sodium formate pH 4.0, 15 mM β-mercaptoethanol, and 1 mM Na EDTA at room temperature. The column was washed with 200 ml equilibration buffer, and the protein eluted with a 1.0 liter linear gradient from 0-0.4 M NaCl. The HIV protein (18 Kd) eluted at approximately 0.2 M NaCl as assayed by SDS-polyacrylamide gel electrophoresis.

   Further purification was obtained by pooling Sub 2-containing fractions and applying to an S-200 (Pharmacia) gel filtration column equilibrated in the same buffer as the previous column.

Example 7 – Synthetic peptides

Synthesis of peptides can be done by a variety of established procedures, for example, automated peptide synthesis. Peptides were assembled by solid-phase synthesis on cross-linked polystyrene beads starting from the carboxyl terminus and adding amino acids in a step-wise fashion (Merrifield, R.B. [1963] S. Am. Chem. Soc. 85:2149). Each synthesis was performed on an automated peptide synthesizer (Applied Biosystems 430-A) using standard t-Boc chemistry. Amino acids were coupled as highly reactive symmetric anhydrides formed immediately prior to use. To minimize

coupling difficulties, dimethylformamide was used as the coupling buffer. The quantitative ninhydrin assay was used to measure the efficiency of coupling after each amino acid addition (Sarin, V.K., S.B.H. Kent, J.P. Tam, R.B. Merrifield [1981] Anal. Biochem. 117:147 1981).

All peptides were deprotected and cleaved from the polystyrene support using an alternative to HF cleavage. Resin containing peptide was resuspended in a mixture of trifluoroacetic acid, trifluoromethane sulfonic acid, and organic thiol scavengers (Tam, J.P., W.F. Heath, R.B. Merrifield [1986] J. Am. Chem. Soc. 108:5242). Soluble peptide was precipitated with ethyl ether and, after removing ether, resuspended in 200 mM sodium carbonate, 3 M guanidine HCl. The crude peptides were purified by reverse-phase chromatography on a 1.0 cm x 25 cm Vidac semi-preparative C_{18} column. The buffers employed were: (A) 0.1% trifluoroacetic acid in H_2O, and (B) 0.1% trifluoroacetic acid in 80% acetonitrile/20% H_2O. Gradient elution was utilized to elute the bound peptide and collected fractions were further analyzed to identify pure product. Peptide identity was confirmed by amino acid analysis following 6 N HCl hydrolysis. The synthesis included the addition of terminal amino acids not homologous to HIV for purposes of labeling, cross-linking, or structure of the peptide. These non-HIV amino acids are indicated in parenthesis.

The product of synthesis can be further purified by a number of established separatory techniques, for example, ion exchange chromatography.

Example 8 – Construction of and expression from plasmid $pPB1_{RF}$

Plasmid $pPB1_{RF}$, which contains approximately 565 bp of DNA encoding essentially the HIV_{RF} env gene from the PvuII site to the BglII site, and from which is synthesized an approximately 27 Kd fusion protein containing this portion of the gp120 envelope protein can be constructed as follows:

1. Synthesizing DNA fragment in Table 4A.
2. Restricting plasmid pREV2.2 with EcoRV and BamHI and isolating the large fragment, approximately 4 Kd, from an agarose gel.
3. Ligating the fragment prepared in 1. with the pREV2.2 fragment in a volume of 20 μl using T4 DNA ligase, transforming the ligation mixture into competent cell strain CAG 629, and selecting ampicillin-resistant transformants.
4. Selecting such transformants, by appropriate restriction patterns, that have the gp120 fragment cloned in the proper orientation to generate a fusion protein. When the strain harboring this recombinant plasmid is grown at 32°C in 2% medium containing 50 μg/ml ampicillin and the total complement of cellular proteins electrophoresed on an SDS-polyacrylamide gel, a protein of approximately 27 Kd can be visualized by

either coomassie blue staining or by Western blot analysis using as probe selected sera from HIV infected individuals.

Example 9 – Purification of recombinant protein containing HIV envelope sequences from plasmid $pPB1_{RF}$

1. Growth of cells: Cells were grown in a 10-liter volume in a Chemap fermentor in 2% medium. Fermentation temperature was 30°C, the pH was 6.8, and air was provided at 1 vvm. Plasmid selection was provided by 20 μg/ml chloramphenicol. Typical cell yield (wet weight) was 30 g/L.

2. Cell lysis: 50 g, wet cell weight, of E. coli containing the recombinant HIV envelope fusion protein were resuspended in a final volume of 100 ml in 50 mM Tris-Cl pH 8.0, 5 mM KEDTA, 5 mM DDT, 15 mM β-mercaptoethanol, 0.5% TRITON™X-100, and 5 mM PMSF. 300 mg lysozyme was added and the suspension incubated for 30 min at room temperature.

   This material was lysed using a BEAD-BEATER™ containing an equal volume of 0.1-0.15 mm glass beads. The lysis was done for 6 min at room temperature in 1 min intervals. The liquid was separated from the beads and centrifuged for 2.5 hr at 20,000 xg. The supernatant was removed and the pellet was resuspended in 100 ml 8 M urea, 20 mM Tris-Cl pH 8.0, 5 mM DTT, 15 mM β-mercaptoethanol, 5 mM PMSF, and 1 mM KEDTA. The pellet was solubilized using a polytron homogenizer and centrifuged at 20,000 xg for 2 hr.

   The supernate (90 ml) was dialysed against 4 liters of 8 M urea, 20 mM HEPES, pH 6.5, 1 mM EDTA, and 15 mM β-mercaptoethanol. Dialysis was done each time for 8 hr or longer with three changes of buffer. Spectrophor dialysis tubing with a 3.5 Kd MW cut-off was used.

3. CM Chromatography: The dialysate was loaded onto a 100 ml column (2.5 cm x 20 cm) packed with CM FAST FLOW SEPHAROSE™ equilibrated in 8 M urea, 10 mM HEPES pH 6.5, 15 mM β-mercaptoethanol, and 1 mM Na EDTA at room temperature. The column was washed with 200 ml equilibrium buffer, and the protein eluted with a 1.0 liter linear gradient from 0-0.4 M NaCl. The HIV protein (26 Kd) eluted at approximately 0.2 M NaCl as assayed by SDS-polyacrylamide gel electrophoresis.

   Further purification was obtained by pooling $PB1_{RF}$-containing fractions and applying to an S-300 gel filtration column equilibrated in the same buffer as the previous column.

Example 10 – Construction of and expression from plasmid $pPB1_{MN}$

Plasmid $pPB1_{MN}$, which contains approximately 600 bp of DNA encoding essentially the HIV_{MN} env gene from the BglII site to the BglII site, and from which is synthesized an approximately 28 Kd fusion protein containing this portion of the gp120 envelope protein, can be constructed as follows:

1. Synthesizing DNA fragment in Table 5A.
2. Restricting plasmid pREV2.2 with BamHI.
3. Ligating the fragment prepared in 1. with the pREV2.2 fragment in a volume of 20 μl using T4 DNA ligase, transforming the ligation mixture into competent cell strain CAG 629, and selecting ampicillin-resistant transformants.
4. Selecting such transformants, by appropriate restriction patterns, that have the gp120 fragment cloned in the proper orientation to generate a fusion protein. When the strain harboring this recombinant plasmid is grown at 32°C in 2% medium containing 50 μg/ml ampicillin and the total complement of cellular proteins electrophoresed on an SDS-polyacrylamide gel, a protein of approximately 28 Kd can be visualized by either coomassie blue staining or by Western blot analysis using as probe selected sera from HIV infected individuals.

Example 11 – Purification of recombinant protein containing HIV envelope sequences from plasmid $pPB1_{MN}$

1. Growth of cells: Cells were grown in a 10-liter volume in a Chemap fermentor in 2% medium. Fermentation temperature was 30°C, the pH was 6.8, and air was provided at 1 vvm. Plasmid selection was provided by 20 μg/ml chloramphenicol. Typical cell yield (wet weight) was 30 g/l.
2. Cell lysis: 50 g, wet cell weight, of E. coli containing the recombinant HIV envelope fusion protein were resuspended in a final volume of 100 ml in 50 mM Tris-Cl pH 8.0, 5 mM KEDTA, 5 mM DTT, 15 mM β-mercaptoethanol, 0.5% TRITON™X-100, and 5 mM PMSF. 300 mg lysozyme was added and the suspension incubated for 30 min at room temperature.

   This material was lysed using a BEAD-BEATER™ containing an equal volume of 0.1-0.15 mm glass beads. The lysis was done for 6 min at room temperature in 1 min intervals. The liquid was separated from the beads and centrifuged for 2.5 hr at 20,000 xg. The supernatant was removed and the pellet was resuspended in 100 ml 8 M urea, 20 mM Tris-Cl pH 8.0, 5 mM DTT, 15 mM β-

mercaptoethanol, 5 mM PMSF, and 1 mM KEDTA. The pellet was solubilized using a polytron homogenizer and centrifuged at 20,000 xg for 2 hr.

The supernate (90 ml) was dialysed against 4 liters of 8 M urea, 20 mM HEPES, pH 6.5, 1 mM EDTA, and 15 mM β-mercaptoethanol. Dialysis was done each time for 8 hr or longer with three changes of buffer. Spectraphor dialysis tubing with a 3.5 Kd MW cut-off was used.

3. CM Chromatography: The dialysate was loaded onto a 100 ml column (2.5 cm x 20 cm) packed with CM FAST FLOW SEPHAROSE™ equilibrated in 8 M urea, 10 mM HEPES pH 6.5, 15 mM β-mercaptoethanol, and 1 mM KEDTA at room temperature. The column was washed with 200 ml equilibration buffer, and the protein eluted with a 1.0 liter linear gradient from 0-0.4 M NaCl. The HIV protein (28 Kd) eluted at approximately 0.2 M NaCl as assayed by SDS-polyacrylamide gel electrophoresis.

Further purification was obtained by pooling $PB1_{MN}$-containing fractions and applying to an S-300 gel filtration column equilibrated in the same buffer as the previous column.

Example 12 – Construction of and expression from plasmid $pPB1_{SC}$

Plasmid $pPB1_{SC}$, which contains approximately 570 bp of DNA encoding essentially the HIV_{SC} env gene from the PvuII site to the BglII site, and from which is synthesized an approximately 26 Kd fusion protein containing this portion of the gp120 envelope protein, can be constructed as follows:

1. Synthesizing DNA fragment in Table 6A.
2. Restricting plasmid pREV2.2 with EcoRV and BamHI and isolating the large fragment, approximately 4 Kd, from the agarose gel.
3. Ligating the fragment prepared in 1. with the pREV2.2 fragment in a volume of 20 μl using T4 DNA ligase, transforming the ligation mixture into competent cell strain CAG 629, and selecting ampicillin-resistant transformants.
4. Selecting such transformants, by appropriate restriction patterns, that have the gp120 fragment cloned in the proper orientation to generate a fusion protein. When the strain harboring this recombinant plasmid is grown at 32°C in 2% medium containing 50 μg/ml ampicillin and the total complement of cellular proteins electrophoresed on an SDS-polyacrylamide gel, a protein of approximately 26 Kd can be visualized by either coomassie blue staining or by Western blot analysis using as probe selected sera from HIV infected individuals.

Example 13 – Purification of recombinant protein containing HIV envelope sequences from plasmid $pPB1_{SC}$

1. Growth of cells: Cells were grown in a 10-liter volume in a Chemap fermentor in 2% medium. Fermentation temperature was 30°C, the pH was 6.8, and air was provided at 1 vvm. Plasmid selection was provided by 20 μg/ml chloramphenicol. Typical cell yield (wet weight) was 30 g/L.

2. Cell lysis: 50 g, wet cell weight, of E. coli containing the recombinant HIV envelope fusion protein were resuspended in a final volume of 100 ml in 50 mM Tris-Cl pH 8.0, 5 mM KEDTA, 5 mM DTT, 15 mM β-mercaptoethanol, 0.5% TRITON™X-100, and 5 mM PMSF. 300 mg lysozyme was added and the suspension incubated for 30 min at room temperature.

   This material was lysed using a BEAD-BEATER™ containing an equal volume of 0.1-0.15 mm glass beads. The lysis was done for 6 min at room temperature in 1 min intervals. The liquid was separated from the beads and centrifuged for 2.5 hr at 20,000 xg. The supernatant was removed and the pellet was resuspended in 100 ml 8 M urea, 20 mM Tris-Cl pH 8.0, 5 mM DTT, 15 mM β-mercaptoethanol, 5 mM PMSF, and 1 mM KEDTA. The pellet was solubilized using a polytron homogenizer and centrifuged at 20,000 xg for 2 hr.

   The supernate (90 ml) was dialysed against 4 liters of 8 M urea, 20 mM HEPES, pH 6.5, 1 mM EDTA, and 15 mM β-mercaptoethanol and 1 mM KEDTA at room temperature. The dialysate was loaded onto a 100 ml column (2.5 cm x 20 cm) packed with CM FAST FLOW SEPHAROSE™ equilibrated in 8 M urea, 10 mM HEPES pH 6.5, 15 mM β-mercaptoethanol, and 1 mM KEDTA at room temperature. The column was washed with 200 ml equilibration buffer, and the protein eluted with a 1.0 liter linear gradient from 0-0.4 M NaCl. The HIV protein (26 Kd) eluted at approximately 0.2 M NaCl as assayed by SDS-polyacrylamide gel electrophoresis.

   Further purification was obtained by pooling $PB1_{SC}$-containing fractions and applying to an S-300 gel filtration column equilibrated in the same buffer as the previous column.

Example 14 – Construction of and expression from plasmid $pPB1_{WMJ2}$

Plasmid $pPB1_{WMJ2}$, which contains approximately 560 bp of DNA encoding essentially the HIV_{WMJ2} env gene and from which is synthesized an approximately 26 Kd fusion protein containing this portion of the gp120 envelope protein, can be constructed as follows:

1. Synthesizing DNA fragment in Table 7A.
2. Restricting plasmid pREV2.2 with EcoRV and BamHI and isolating the large fragment, approximately 4 Kd, from an agarose gel.
3. Ligating the fragment prepared in 1. with the pREV2.2 fragment in a volume of 20 μl using T4 DNA ligase, transforming the ligation mixture into competent cell strain CAG 629, and selecting ampicillin-resistant transformants.
4. Selecting such transformants, by appropriate restriction patterns, that have the gp120 fragment cloned in the proper orientation to generate a fusion protein. When the strain harboring this recombinant plasmid is grown at 32°C in 2% medium containing 50 μg/ml ampicillin and the total complement of cellular proteins electrophoresed on an SDS-polyacrylamide gel, a protein of approximately 26 Kd can be visualized by either coomassie blue staining or by Western blot analysis using as probe selected sera from HIV infected individuals.

Example 15 – Purification of recombinant protein containing HIV envelope sequences from plasmid $pPB1_{WMJ2}$

1. Growth of cells: Cells were grown in a 10-liter volume in a Chemap fermentor in 2% medium. Fermentation temperature was 30°C, the pH was 6.8, and air was provided at 1 vvm. Plasmid selection was provided by 20 μg/ml chloramphenicol. Typical cell yield (wet weight) was 30 g/l.
2. Cell lysis: 50 g, wet cell weight, of E. coli containing the recombinant HIV envelope fusion protein were resuspended in a final volume of 100 ml in 50 mM Tris-Cl pH 8.0, 5 mM KEDTA, 5 mM DTT, 15 mM β-mercaptoethanol, 0.5% TRITON™X-100, and 5 mM PMSF. 300 mg lysozyme was added and the suspension incubated for 30 min at room temperature.

   This material was lysed using a BEAD-BEATER™ containing an equal volume of 0.1-0.15 mm glass beads. The lysis was done for 6 min at room temperature in 1 min intervals. The liquid was separated from the beads and centrifuged for 2.5 hr at 20,000 xg. The supernatant was removed and the pellet was resuspended in 100 ml 8 M urea, 20 mM Tris-Cl pH 8.0, 5 mM DTT, 15 mM β-mercaptoethanol, 5 mM PMSF, and 1 mM KEDTA. The pellet was solubilized using a polytron homogenizer and centrifuged at 20,000 xg for 2 hr.

   The supernate (90 ml) was dialysed against 4 liters of 8 M urea, 20 mM HEPES, pH 6.5, 1 mM EDTA, and 15 mM β-mercaptoethanol. Dialysis was done

each time for 8 hr or longer with three changes of buffer. Spectraphor dialysis tubing with a 3.5 Kd MW cut-off was used.

3. CM Chromatography: The dialysate was loaded onto a 100 ml column (2.5 cm x 20 cm) packed with CM FAST FLOW SEPHAROSE™ equilibrated in 8 M urea, 10 mM HEPES pH 6.5, 15 mM β-mercaptoethanol, and 1 mM Na EDTA at room temperature. The column was washed with 200 ml equilibration buffer, and the protein eluted with a 1.0 liter linear gradient from 0-0.4 M NaCl. The HIV protein (26 Kd) eluted at approximately 0.2 M NaCl as assayed by SDS-polyacrylamide gel electrophoresis.

Further purification was obtained by pooling $PB1_{WMJ2}$-containing fractions and applying to an S-300 gel filtration column equilibrated in the same buffer as the previous column.

Example 16 – Cell fusion inhibition

The in vitro fusion of HIV infected cells with T4 positive T cells is measured in the presence and absence of immune serum. This is a well known assay (Putney et al. [1986] Science 234:1392-1395).

Chronically infected cells and uninfected cells (1:15) are mixed and incubated 24 hr. Foci of fused cells (giant cells) are then counted (usually about 60). Dilution of an immune serum, for example, serum to the entire HIV envelope (gp160) or to a protein or peptide of the invention, is added when the cells are mixed. A 90% decrease in giant cells after 24 hr indicates the immune serum can block fusion. This assay can be done with cells infected with various virus strains, for example, HIV_B and HIV_{RF}.

Example 17 – Competition cell fusion

Using the assay described in Example 16, one can determine if proteins or peptides contain the epitope recognized by antibodies that are responsible for cell fusion inhibition. For example, fusion inhibition of HIV_{IIIB} infected cells by antiserum to the PB1-III_B protein of the parent application is abated by addition of PB1-III_B protein to 5 μg/ml. Using antiserum to PB1-III_B and adding any one of the proteins or peptides, for example, Sub 2, Sub 1, CNBr1, peptide 135 or peptide 136 at 5 μg/ml totally blocks the activity of the PB1 antiserum. Additionally, antiserum to PB1-RF that is capable of neutralizing HIV-RF can be blocked in this activity by peptide 139. A peptide containing only the central portion of the peptide 139, e.g., peptide 339, also can block the fusion inhibition activity of antiserum to PB1-RF. This, for the first time, localizes the critical amino acids

necessary to elicit neutralization or block fusion inhibiting antibody to a ten amino acid sequence (e.g., peptide 339).

Example 18 – Co-immunization of PB1-III_B and $PB1_{RF}$

Antisera from an animal immunized with two PB1 proteins from HIV_{IIIB} and HIV_{RF} isolates were capable of blocking cell fusion of both HIV_{IIIB}- and HIV_{RF}-infected cells. This demonstrates that co-immunization with separate proteins containing envelope sequences of two HIV isolates elicits an immune response capable of neutralizing both isolates. This novel property of small proteins or peptides blocking immune serum has not been described before.

Some of the proteins and peptides of the subject invention contain the entire epitope for raising humoral immune responses that neutralize HIV infection and block HIV infected cell fusion. This is shown by these proteins and peptides competing these activities out of serum from animals immunized with the entire HIV envelope. More specifically, proteins and peptides that can compete the activities from anti-gp160 or anti-PB1 sera are Sub 2, Sub 1, CNBr1, peptide 135, and peptide 136.

The proteins and peptides of the invention also can be used to stimulate a lymphocyte proliferative response in HIV infected humans. This then would stimulate the immune system to respond to HIV in such individuals.

Example 19 – Construction of and expression from plasmid $pPB1_{IIIB}$

Plasmid pPB1, which contains approximately 540 bp of DNA encoding essentially the HIV env gene from the PvuII site to the BglII site, and from which is synthesized an approximately 26 Kd fusion protein containing this portion of the gp120 envelope protein, can be constructed as follows:

1. Synthesizing the DNA with the sequence shown in Table 8: This DNA fragment can be synthesized by standard methods and encodes a portion of gp120. It has a blunt end on the 5' end and an end which will ligate with a BamHI overhang on the 3' end.
2. Restricting 5 μg plasmid pREV2.2 with EcoRV and BamHI and isolating the large fragment, approximately 4 Kd, from an agarose gel.
3. Ligating 0.1 μg of the fragment in Table 8 with 0.1 μg of the pREV2.2 fragment in a volume of 20 μl using T4 DNA ligase, transforming the ligation mixture into competent cell strain SG20251, and selecting ampicillin-resistant transformants.
4. Using the AhaIII restriction pattern of purified plasmid, selecting cells harboring the recombinant plasmid with the synthesized fragment in the orientation whereby the fragment blunt end ligated to the REV2.2 EcoRV end and the BamHI overhanging ends ligated together. AhaIII digestion of the proper plasmid gives fragment lengths of approximately 1210, 1020, 750, 690, 500, 340, and 20 base pairs. When the strain

harboring this recombinant plasmid is grown in 2% medium containing 50 μg/ml ampicillin and the total complement of cellular proteins electrophoresed on an SDS-polyacrylamide gel, a protein of approximately 26 Kd can be visualized by either coomassie blue staining or by Western blot analysis using as probe selected sera from AIDS, ARC, or HIV infected individuals.

Example 20 – Purification of recombinant protein containing HIV envelope sequences from plasmid $pPB1_{IIIB}$

1. Growth of cells: Cells were grown in a 10-liter volume in a Chemap fermentor in 2% medium. Fermentation temperature was 37°C, the pH was 6.8, and air was provided at 1 vvm. Plasmid selection was provided by 50 μg/ml ampicillin and 20 μg/ml chloramphenicol. Typical cell yield (wet weight) was 30 g/l.

2. Cell lysis: 50 g, wet cell weight, of E. coli containing the recombinant HIV envelope fusion protein were resuspended in a final volume of 100 ml in 50 mM Tris-Cl pH 8.0, 5 mM KEDTA, 5 mM DTT, 15 mM β-mercaptoethanol, 0.5% TRITON™X-100, and 5 mM PMSF. 300 mg lysozyme was added and the suspension incubated for 30 min at room temperature.

   This material was lysed using a BEAD-BEATER™ (Biospec Products, Bartlesville, OK) containing an equal volume of 0.1-0.15 mm glass beads. The lysis was done for 6 min at room temperature in 1 min intervals. The liquid was separated from the beads and centrifuged for 2.5 hr at 20,000 xg. The supernatant was removed and the pellet was resuspended in 100 ml 6 M guanidine-hydrochloride, 20 mM Tris-Cl pH 8.0, 5 mM DTT, 15 mM β-mercaptoethanol, 5 mM PMSF, and 1 mM KEDTA. The pellet was solubilized using a polytron homogenizer and centrifuged at 20,000 xg for 2 hr.

   The supernate (90 ml) was dialysed against 4 liters of 8 M urea, 20 mM Tris-Cl, pH 8.0, 1 mM EDTA, and 15 mM β-mercaptoethanol. Dialysis was done each time for 8 hr or longer with three changes of buffer. Spectraphor dialysis tubing (S/P, McGraw Park, IL) with a 3.5 Kd MW cut-off was used.

3. CM Chromatography: The dialysate was loaded onto a 550 ml column (5 cm x 28 cm) packed with CM FAST FLOW SEPHAROSE™ equilibrated in 8 M urea, 10 mM potassium phosphate pH 7.0, 15 mM β-mercaptoethanol, and 1 mM KEDTA at room temperature. The column was washed with 2 liters equilibration buffer, and the protein eluted with a 5.0 liter linear gradient from 0-0.4 M NaCl. The HIV

protein (26 Kd) eluted at approximately 0.2 M NaCl as assayed by SDS-polyacrylamide gel electrophoresis.

Example 21 – Immunization with two or more peptides to obtain broadly neutralizing antisera

Five peptides, i.e., peptide 135, peptide 139, peptide 141, peptide 142 and peptide 143, were cross-linked individually to carrier proteins and used to immunize goats. Each peptide is capable of eliciting type specific neutralization when used individually as an immunogen. Synthetic peptides were cross-linked through a sulfhydryl bond to keyhole limpet hemocyanin (KLH) by using Succinimidyl-4-(n-Maleimidomethyl)Cyclohexane 1-Carboxylate (Pierce). The ratio of peptide to KLH was 1:2 by weight. 200 μg of each cross-linked peptide was used in the immunization cocktail (a total of 1 mg of 5 peptides, 2 mg of KLH). This method of crosslinking or immunization regimen is but an example and not meant to be limiting. After four immunizations, immune sera was tested for neutralization of these five HIV isolates as well as distinctly different isolates. The immune serum could block fusion of cells infected with any of five isolates from which the peptide sequences were derived. In addition, the serum neutralized other variants not used in the immunization.

Equivalent broad neutralizing sera may also be obtained by variations of this immunogen. For example, using more than five peptides having the amino acid sequence derived from the principal neutralizing domain from more than five variants. Alternatively, a single peptide (e.g., peptide 64 or peptide 74) containing segments homologous to diverse HIV variants may also be used to elicit broad neutralizing antibody.

Example 22 – Sequential Immunization with Two or More Peptides as a Method to Elicit Broad Neutralizing Antisera

An immunization protocol capable of eliciting broad neutralizing antibodies may take the form of initial immunization with a peptide or protein antigenically equivalent to the principal neutralizing domain, or segments thereof. The initial immunization is followed with a second immunization. The initial immunization could be done with, for example, peptide 135, peptide 139, peptide 141, peptide 142, or peptide 143, with subsequent immunization with, for example, one or more of the following peptides:

RP57	Ile	Asn	Cys	Thr	Arg	Pro	Ala	His	Cys	Asn	Ile	Sr
RP55	Ala	His	Cys	Asn	Ile	Ser						
RP75	(Ala	Ala	Ala	Ala	Ala	Ala)	Gly	Pro	Gly	Arg	(Ala	
	Ala	Ala	Ala	Ala	Ala	Cys)						
RP56	Ile	Asn	Cys	Thr	Arg	Pro						
RP59	Ile	Gly	Asp	Ile	Arg	Gln	Ala	His	Cys	Asn	Ile	Sr

The method is to immunize with a protein or peptide and then boost the immune response to a defined subset of the original immunogen. This immunization method may be useful in vaccine methodology and also to generate broad neutralizing polyclonal or monoclonal antibodies for therapeutic applications.

Example 23 – Identification of Critical Segments of the Principal Neutralizing Domain

Certain segments of the principal neutralizing domain have been found to be capable of eliciting the antigenic and immunogenic responses which are associated with the principal neutralizing domain as a whole. For example, a region of the principal neutralizing domain known as the "tip of the loop" has been shown to be capable of raising, and/or binding with, neutralizing antibodies. This capability is observed for the "tip of the loop" of a variety of HIV variants.

The tip of the loop comprises a three amino acid segment which is highly conserved between HIV variants, together with various amino acids which occur on either side of the three conserved amino acids. The three conserved amino acids, which are Gly Pro Gly, usually occur at, or about, positions 311, 312, and 313 of the HIV envelope protein.

The "tip of the loop" comprises the Gly Pro Gly segment together with the 2 to 8 amino acids which flank either or both sides of this segment in any given HIV variant. The amino acids which flank the conserved segment may be any of the 20 natural amino acids, in any sequential order. Although the amino acid sequence of the principal neutralizing domain varies between different HIV-1 isolates, conservation at particular positions, for example at the tip of the loop, suggests that certain amino acids at these positions are necessary for virus function.

Example 24 – Sequence of the Principal Neutralizing Determinant from Randomly Selected HIV-1 Isolates

Sequences of the principal neutralizing domains (PNDs) from random field isolates were obtained in order to determine the degree of heterogeneity within this region of the envelope protein. Peripheral blood lymphocytes (PBLs) from randomly selected HIV-1 infected donors were either co-cultured with uninfected PBLs or the virus isolates were adapted to CD4 cell lines. DNA was extracted from these infected cells and a 240 base pair region encoding the PND was amplified by polymerase chain reaction using oligonucleotide primers that hybridize with flanking conserved regions. This product was cloned into pUC19 and the sequence of the PND from one or more clones from each original isolate were determined. Because of the heterogeneity of the virus population within one infected individual, when two or more sequences were obtained from one PCR reaction, these sequences sometimes differed.

The data obtained from nearly 100 individuals (some infected with a heterogeneous virus population) was evaluated along with previously obtained HIV sequence information. Table 9 lists 138 PND sequences from HIV isolates. These sequences indicate that, despite the well-known and frequently cited variability in the amino acid sequence of the HIV envelope protein, there is actually a high degree of conservation in the immunologically critical PND region, particularly in the region at the center of the PND. Specifically, the Gly-Pro-Gly sequence at the "tip of the loop" occurs in over 90% of the variants. Furthermore, other amino acids at certain positions on each side of the G-P-G were also found to be highly conserved. Table 10 shows the frequency of occurrence of the various amino acids at each position in the PND. In addition to the very strong conservation of the glycines flanking the central proline, there is strong conservation at several other positions (e.g., R at x_{12}, P at x_{11}, G at y_{11}, R at y_{14}, and A at y_{16}).

A comparison of the relative frequency of variations of a 17-amino acid segment centered about the G-P-G sequence is shown in Table 11. In this table, the sequence which reflects the most commonly occurring amino acids at each position is listed first. The dashes indicate identity with the consensus sequence. The remaining sequences are ordered from 2 to 138, according to their homology to the consensus sequence. Thus, the sequences at the top of the table display the greatest homology with the consensus sequence. Sequences far down the table display less homology. For example, the amino acid sequences from isolates IIIB and LAV-BRU occur at positions 92 and 93, respectively, on this table. This indicates that these isolates have only limited homology with the consensus sequence. The present research shows that HIV viruses such as IIIB and LAV-BRU having the Gln-Arg (Q-R) dipeptide to the left of the Gly-Pro-Gly sequence are relatively uncommon. By contrast, the MN-like sequence in this region (...I H I G P G...) is the most common. The present research shows that principal neutralizing domains of other commonly studied variants comprise relatively uncommon sequences.

Although the subject invention pertains to the discovery of certain highly conserved regions in the principal neutralizing domain, there remains some degree of variability in this region among the various isolates. This variability includes "missing" or "added" amino acids at certain points in the sequence. Of course, "missing" or "added" amino acids can cause difficulty in devising aesthetically pleasing tables showing the sequences. However, these missing or added amino acids pose no difficulty to a person skilled in the art in terms of locating the highly conserved regions which are critical to the subject invention. Table 9 shows one representation of 138 PND sequences. Table 11 uses a slightly different representation to show the same PND sequences. The primary discussion of sequence conservation can probably best be visualized by reference to the representation shown in Table 11. However, it should be noted that the existence of more than one means for representing these

sequences does not compromise the ability of the skilled artisan to accurately locate the sequences or the conserved regions.

With the discovery of commonly occurring amino acid sequences, it is possible, for the first time, to develop prophylactic and therapeutic compositions which can predictably elicit and/or bind with neutralizing antibodies to a broad range of HIV variants. The generalized formula for such a composition can be as follows:

a x G z G y b

wherein x is 0 to 13 amino acids in length;

y is 0 to 17 amino acids in length; and

z is P, A, S, Q, or L; and

either a or b, but not both, may be omitted; either a or b individually may comprise any one of the following: cysteine, a protein or other moiety capable of enhancing immunogenicity, a peptide from an HIV principal neutralizing domain, a peptide capable of stimulating T-cells, or a general immune stimulant.

Further analysis of common sequence patterns reveals certain specific patterns which are very common among HIV isolates. Examples of these common sequences are shown in Table 12 and Figure 1. These common patterns, which are known only as a result of the research and discoveries described here, can be used to make pharmaceutical and diagnostic compositions which can be used with a broad range of HIV isolates. This, of course, can be of critical importance, given the large number of HIV variants which are now known to exist.

Table 12 is a compilation of the common sequence patterns that occur in the region at the tip of the loop. For example, approximately 60% of the HIV isolates contain the core sequence I a I G P G R (a represents several different residues), approximately 50% contain the sequence I G P G R A, and approximately 40% contain G P G R A F. When a His residue is present at the a position in I a I G P G R, this sequence occurs in approximately 30% of the HIV isolates. A vaccine composition comprising a mixture of peptides having the sequence I a I G P G R where all of the possible replacements for the a are present, is capable of eliciting antibodies which neutralizes a majority of HIV variants. Preferably, for use as immunogens, the peptides are linked to carrier proteins or adjuvants as described in Example 21.

As shown in Figure 1, common sequence patterns are also apparent within the 17 amino acid segment. Sequences which were isolated 4 or more times are highlighted. These commonly occurring sequences can be used to formulate vaccine cocktails which elicit a broadly neutralizing response. For example, a potential cocktail might contain peptides from each of the eight groups represented. Alternatively, the peptide sequences may be presented as a hybrid polypeptide containing the principal

neutralizing domain from two or more of these groups. Preferably, such a cocktail will contain peptides which will be capable of raising antibodies which neutralize at least 70% and most preferably at least 90% of HIV variants.

The antigens of the subject invention can be identified by their ability to raise antibodies which bind to certain amino acid sequences. For example, particularly advantageous antigenic compounds would raise antibodies which bind to common amino acid sequences such as G-P-G-R-A-F, I-G-P-G-R-A-F, I-G-P-G-R-A, I-a-I-G-P-G-R, I-a-I-G-P-G-R-A, and I-a-I-G-P-G-R-A-F, where a is any of the 20 amino acids.

From Table 10 it can be seen that a polypeptide representing the occurrence of amino acids in all of the variants can be represented as follows:

$x_{13}\ x_{12}\ x_{11}\ x_{10}\ x_9\ x_8\ x_7\ x_6\ x_5\ x_4\ x_3\ x_2\ x_1\ G\ z\ G\ y_2\ y_3\ y_4\ y_5\ y_6\ y_7\ y_8\ y_9\ y_{10}\ y_{11}\ y_{12}\ y_{13}\ y_{14}\ y_{15}\ y_{16}\ y_{17}$, wherein

x_1 is I, R, M, IQR, V, L, K, F, S, G, Y, SRG, or YQR;
x_2 is H, R, Y, T, S, P, F, N, A, K, G, or V;
x_3 is I, L, M, T, V, E, G, F, or Y;
x_4 is R, S, G, H, A, K, or not present;
x_5 is K, R, I, N, Q, A, IR, RQ, or not present;
x_6 is R, K, S, I, P, Q, E, G, or T;
x_7 is T, K, V, I, A, R, P, or E;
x_8 is N, NV, Y, KI, I, T, DK, H, or K;
x_9 is N, S, K, E, Y, D, I, or Q;
x_{10} is N, Y, S, D, G, or H;
x_{11} is P;
x_{12} is R, I, or K;
x_{13} is T, I, M or A;
z is P, A, Q, S, or L;
y_1 is R, K, Q, G, S, or T;
y_2 is A, V, N, R, K, T, S, F, P, or W;
y_3 is F, I, V, L, W, Y, G, S, or T;
y_4 is Y, V, H, L, F, S, I, T, M, R, VH, or FT;
y_5 is T, A, V, Q, H, I, S, Y, or not present;
y_6 is T, R, I, Q, A, M, or not present;

y_7 is G, E, K, R, T, D, Q, A, H, N, P, or not present;
y_8 is R, Q, E, K, D, N, A, G, S, I, or not present;
y_9 is I, V, R, N, G, or not present;
y_{10} is I, T, V, K, M, R, L, S, E, Q, A, or not present;
y_{11} is G, R, E, K, H, or not present;
y_{12} is D, N, I, R, T, S, or not present;
y_{13} is I, M, ME, L, or not present;
y_{14} is R, G, K, S, E, or not present;
y_{15} is Q, K, or R;
y_{16} is A; and
y_{17} is H, Y, R, or Q.

Monoclonal and/or polyclonal antibodies with broad neutralizing activity can be generated using the commonly occurring peptide sequences for use in prophylactic or therapeutic compositions. The commonly occurring sequences described here can be used in much the same way as the other peptides described in this application. For example, these peptides can be modified in order to provide T-lymphocyte stimulation, general immune stimulation, to enhance immunogenicity or solubility, or to reduce toxicity. The peptides may also be modified by addition of terminal cysteine residue(s) or by conjugation to a carrier protein, adjuvant, spacer, and/or linker. The peptides may be fused with other HIV epitopes to produce a multiepitope polypeptide which could be useful with an even greater number of HIV variants. Also, the peptides can be circularized by bonding between cysteine residues. The cysteine residues used to make such circularized peptides could be the naturally occurring cysteine residues at the ends of the principal neutralizing domain, or cysteine residues may be added to the terminal ends of the peptides.

Additionally, vaccine compositions may include peptides containing T helper cell epitopes in combination with protein fragments containing the principal neutralizing domain. Several of these epitopes have been mapped within the HIV envelope, and these regions have been shown to stimulate proliferation and lymphokine release from lymphocytes. Providing both of these epitopes in a vaccine composition may result in the stimulation of both humoral and cellular immune responses.

Example 25 – Construction and Cloning of Multi-Epitope Genes

Synthetic genes can be constructed which encode proteins comprised of the neutralizing epitopes from more than one HIV isolate. The synthetic gene exemplified here comprises a tandem arrangement of DNA sequences encoding neutralizing epitopes from HIV isolates IIIB, RF, SC, MN, and WMJ1. Each epitope-encoding domain within the gene was designed to encode the 11 amino

acids centered at the common Gly-Pro-Gly sequence at the tip-of-the-loop for each of the isolates. Thus, the multi-epitope gene contained 5 different coding regions, each of which encoded a neutralizing epitope from a different isolate. For this particular construction, the epitope which was chosen for each of the 5 isolates consisted of the Gly-Pro-Gly sequence along with the 4 amino acids on either side of the Gly-Pro-Gly sequence from each of the 5 isolates. Domains coding for other neutralizing epitopes from these isolates could have been incorporated into the multi-epitope gene. Also, genes coding for neutralizing epitopes from other isolates can be used.

The genes were constructed such that the domains were linked by DNA sequences encoding four glycine residues. The composition or length of the linking sequence can be varied but preferably it is a sequence that is non-immunogenic itself. The DNA sequence of the synthetic gene described here was designed such that restriction sites were encoded at either end of the fragment to facilitate cloning into the vector or, alternatively, to permit the construction of longer multi-epitope genes by attachment of 2 or more shorter genes (Figure 2). In addition, a methionine residue was encoded at the 5' end of the gene to facilitate cleavage when produced as part of a fusion protein.

Figure 3 depicts the steps in the construction of the multi-epitope gene described here. The amino acid sequence encoded by this gene is shown in Table 13. The portions of this amino acid sequence which correspond to each of the 5 isolates are identified in Table 13.

Double-stranded subfragments of the full-length gene were first constructed starting with single-stranded synthetic oligomers designed to encode tandem neutralizing epitopes and linking amino acid sequences. Any number of subfragments can be used. In this experiment the gene was divided into two portions, but three, four, or more portions can be used. Four single-stranded oligomers of between 67 and 78 nucleotides in length were synthesized (HEO-1, HEO-2, HEO-3, and HEO-4) (Figure 3). The oligomers were designed in pairs (HEO-1 and 2; HEO-3 and 4) as opposite and adjacent strands of the double-stranded subfragments having 10 (HEO-1 + 2) or 11 (HEO-3 +4) bases of complementary overlap. The oligomers of each pair were mixed and heated 65°C for 5 minutes, then incubated at 37°C for 1 hour to anneal.

After annealing, the complementary strands of each pair were completed ("filled-in") using Sequenase (U.S. Biochem) and the four deoxynucleotide triphosphates. This reaction was incubated for 1 hour at room temperature, heated at 65°C for 3 minutes, and then incubated for an additional hour at 37°C with fresh Sequenase. Double-stranded fragments of 141 (HEO-1+2) and 126 (HEO-3+4) base pairs were generated representing adjacent subfragments of the multi-epitope gene. HEO-1+2 comprised the coding sequences for 3 epitopes plus adjacent linker amino acids; HEO-3+4 extended from the fourth epitope to the end of the gene.

Following the fill-in reaction, the samples were extracted with phenol/chloroform and precipitated with ethanol by standard procedures. The resulting double-stranded DNAs were digested

with HindIII (HEO-1+2) or SacI (HEO-3+4) and purified on a 3% NuSieve agarose gel. The purified fragments were ligated with HindIII + SacI digested pUC19 (New England Biolabs) in a 3-component ligation and transformed into E. coli JM105 cells. The presence of a 256 base pair fragment in pUC19, encoding the full-length multi-epitope gene, was confirmed by restriction analysis and DNA sequencing. The resulting plasmid was designated pUC/MEP-1.

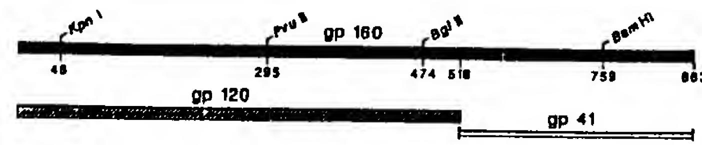
The MEP-1 insert was removed from pUC/MEP-1 and recloned into HindIII + SacI digested pRev2.1 for high-level expression of a fusion protein comprised of a leader portion from the E. coli BG gene fused to the multi-epitope protein. The resulting plasmid, designated pMEP-1-8342, was transformed into E. coli strain SG20251 and the 12.9 Kd multi-epitope fusion protein was identified by coomassie blue staining or Western blot analysis using a probe selected from antisera to the loop-tip peptides from each of the 5 HIV isolates. The fusion protein can be used intact or, alternatively, the leader portion can be cleaved off by cyanogen bromide which cleaves on the carboxy-terminal side of methionine residues. The amino acid sequence of the fusion protein is shown in Table 13A.

The multiepitope peptide can be purified from recombinant cells by methods described above.

Other synthetic genes can be constructed which encode tandem neutralizing epitopes from any number of different HIV isolates using the procedure described above. In addition, variations on the above procedure can be made which are meant to be included in the present invention. For example, the lengths of the neutralizing epitopes encoded by a gene can vary, and there can be variation in the length of the individual epitopes within a single gene. Further, the number of neutralizing epitopes within a multi-epitope gene can vary, and the composition or the length of the amino acid sequences of the epitopes or the linking sequences can be varied from the example that is described herein.

It should be understood that the examples and embodiments described herein are for illustrative purposes only and that various modifications or changes in light thereof will be suggested to persons skilled in the art and are to be included within the spirit and purview of this application and the scope of the appended claims.

This work was supported in part under contract number N01-AI-62558, awarded to Repligen Corporation by the National Institute of Allergy and Infectious Diseases (NIAID).



III_B (BH10)	C T R P N N N T R K S I R I Q R G P G R A F V T I G K I G N M R Q A H C	III_B (BH10)
RF	C - - - - - - - - - S - T K - - - - V I Y A T - Q I - - D I - K - - C	RF
MN	C - - - - Y - K - - - - H I - - - - - - - Y - T K N I - - T I - - - - - C	MN
SC	C - - - - - - - - T R S - H I - - - - - - - Y A T - D I - - D I - - - - - C	SC
WMJ-2	C - - - Y - - V - R S L S I - - - - - - - R - R E I - - I I - - - - - C	WMJ-2
LAV-MAL	C - - - G - - - - - R G - H F - - - Q - L Y - T - I V - D I - R - Y C	LAV-MAL
SF-2	C - - - - - - - - - S - Y I - - - - - - - H - T - R I - - D I - K - - C	SF-2
NY5	C - - - - - - - - K - G - A I - - - - T L Y A R E - I - - D I - - - - - C	NY5
Z3	C - - - G SDKKI - Q S - R I - - - K V - Y A K - G I T - - - - C	Z3
WMJ1	C - - - - - - - V - R R H - E I - - - - - - - Y - G E I R - - I - - - - - C	WMJ1
WMJ3	C - - - - D I A - R - - H I - - - - - - - Y - G K I - - - I - - - - - C	WMJ3
Z6	C - - - Y K - - - Q S T P I - L - Q - L Y -TRGRT- I - - - - - C	Z6
LAVELI	C A - - Y Q - - - Q - T P I - L - Q S L Y - TRSRS I - - - - - C	LAVELI
CDC451	C - - - - - E - - - - V T L - - - - V W Y - T - E I L - - I - - - - - C	CDC451
CDC42	C - - - - - - - - - - V T L - - - - V W Y - T - E I L - - I - - - - - C	CDC42
BAL	C - - - - - - - - - S - H I - - - - - - - Y - T - E I - - D I - - - - - C	BAL
HIV-2	C K - - G - K - V - Q - M L M S - H V - H S H Y Q PIN K R P - - - W C	HIV-2

Table 1.

TABLE 2

LeuAsnGlnSerValGluIleAsnCysThrArgProAsnAsnAsnThrArgLys
SerIleArgIleGlnArgGlyProGlyArgAlaPheValThrIleGlyLysIle
GlyAsnMetArgGlnAlaHisCysAsnIleSerArgAlaLysTrpAsnAsnThr

TABLE 2A

CTGAACCAATCTGTAGAAATTAATTGTACAAGACCCAACAACAATACAAGAAAA
AGTATCCGTATCCAGAGAGGACCAGGGAGAGCATTTGTTACAATAGGAAAAATA
GGAAATATGAGACAAGCACATTGTAACATTAGTAGAGCAAAATGGAATAACACTTT

TABLE 2B

MetLeuArgProValGluThrProThrArgGluIleLysLysLeuAspGlyLeu
TrpAlaPheSerLeuAspArgGluArgValAlaAspLeuAsnGlnSerValGlu
IleAsnCysThrArgProAsnAsnAsnThrArgLysSerIleArgIleGlnArg
GlyProGlyArgAlaPheValThrIleGlyLysIleGlyAsnMetArgGlnAla
HisCysAsnIleSerArgAlaLysTrpAsnAsnThrLeuGlyAlaArgIleLeu
GluAspGluArgAlaSer

TABLE 2C

ATGTTACGTCCTGTAGAAACCCCAACCCGTGAAATCAAAAAACTGGACGGCCTG
TGGGCATTCAGTCTGGATCGCGAACGCGTGGCCGATCTGAACCAATCTGTAGAA
ATTAATTGTACAAGACCCAACAACAATACAAGAAAAAGTATCCGTATCCAGAGA
GGACCAGGGAGAGCATTTGTTACAATAGGAAAAATAGGAAATATGAGACAAGCA
CATTGTAACATTAGTAGAGCAAAATGGAATAACACTTTGGGAGCTCGAATTCTT
GAAGACGAAAGGGCCTCG

TABLE 3

LeuAsnGlnSerValGluIleAsnCysThrArgProAsnAsnAsnThrArgLys
SerIleArgIleGlnArgGlyProGlyArgAlaPheValThrIleGlyLysIle
GlyAsnMetArgGlnAlaHisCysAsnIleSerArgAlaLysTrpAsnAsnThr
LeuLysGlnIleAspSerLysLeuArgGluGlnPheGlyAsnAsnLysThrIle
IlePheLysGlnSerSerGlyGlyAspProGluIleValThrHisSerPheAsn
CysGlyGlyGluPhePheTyrCysAsnSerThrGlnLeuPheAsnSer

TABLE 3A

CTGAACCAATCTGTAGAAATTAATTGTACAAGACCCAACAACAATACAAGAAAA
AGTATCCGTATCCAGAGAGGACCAGGGAGAGCATTTGTTACAATAGGAAAAATA
GGAAATATGAGACAAGCACATTGTAACATTAGTAGAGCAAAATGGAATAACACT
TTAAAACAGATAGATAGCAAATTAAGAGAACAATTTGGAAATAATAAAACAATA
ATCTTTAAGCAGTCCTCAGGAGGGGACCCAGAAATTGTAACGCACAGTTTTAAT
TGTGGAGGGGAATTTTTCTACTGTAATTCAACACAACTGTTTAATAGT

TABLE 3B

MetLeuArgProValGluThrProThrArgGluIleLysLysLeuAspGlyLeu
TrpAlaPheSerLeuAspArgGluArgValAlaAspLeuAsnGlnSerValGlu
IleAsnCysThrArgProAsnAsnAsnThrArgLysSerIleArgIleGlnArg
GlyProGlyArgAlaPheValThrIleGlyLysIleGlyAsnMetArgGlnAla
HisCysAsnIleSerArgAlaLysTrpAsnAsnThrLeuLysGlnIleAspSer
LysLeuArgGluGlnPheGlyAsnAsnLysThrIleIlePheLysGlnSerSer
GlyGlyAspProGluIleValThrHisSerPheAsnCysGlyGlyGluPhePhe
TyrCysAsnSerThrGlnLeuPheAsnSerGlySerSerAsnSer

TABLE 3C

ATGTTACGTCCTGTAGAAACCCCAACCCGTGAAATCAAAAAACTGGACGGCCTG
TGGGCATTCAGTCTGGATCGCGAACGCGTGGCCGATCTGAACCAATCTGTAGAA
ATTAATTGTACAAGACCCAACAACAATACAAGAAAAAGTATCCGTATCCAGAGA
GGACCAGGGAGAGCATTTGTTACAATAGGAAAAATAGGAAATATGAGACAAGCA
CATTGTAACATTAGTAGAGCAAAATGGAATAACACTTTAAAACAGATAGATAGC
AAATTAAGAGAACAATTTGGAAATAATAAAACAATAATCTTTAAGCAGTCCTCA
GGAGGGGACCCAGAAATTGTAACGCACAGTTTTAATTGTGGAGGGGAATTTTTC
TACTGTAATTCAACACAACTGTTTAATAGTGGGAGCTCGAATTCT

TABLE 4

LeuAsnAlaSerValGlnIleAsnCysThrArgProAsnAsnAsnThrArgLys
SerIleThrLysGlyProGlyArgValIleTyrAlaThrGlyGlnIleIleGly
AspIleArgLysAlaHisCysAsnLeuSerArgAlaGlnTrpAsnAsnThrLeu
LysGlnValValThrLysLeuArgGluGlnPheAspAsnLysThrIleValPhe
ThrSerSerSerGlyGlyAspProGluIleValLeuHisSerPheAsnCysGly
GlyGluPhePheTyrCysAsnThrThrGlnLeuPheAsnSerThrTrpAsnSer
ThrGluGlySerAsnAsnThrGlyGlyAsnAspThrIleThrLeuProCysArg
IleLysGlnIleValAsnMetTrpGlnGluValGlyLysAlaMetTyrAlaPro
ProIleSerGlyGlnIleLysCysIleSerAsnIleThrGlyLeuLeuLeuThr
ArgAspGlyGlyGluAspThrThrAsnThrThr

TABLE 4A

CTGAATGCATCTGTACAAATTAATTGTACAAGACCCAACAACAATACAAGAAAA
GACTTACGTAGACATGTTTAATTAACATGTTCTGGGTTGTTGTTATGTTCTTTT

AGTATAACTAAGGGACCAGGGAGAGTAATTTATGCAACAGGACAAATAATAGGA
TCATATTGATTCCCTGGTCCCTCTCATTAAATACGTTGTCCTGTTTATTATCCT

GATATAAGAAAAGCACATTGTAACCTTAGTAGAGCACAATGGAATAACACTTTA
CTATATTCTTTTCGTGTAACATTGGAATCATCTCGTGTTACCTTATTGTGAAAT

AAACAGGTAGTTACAAAATTAAGAGAACAATTTGACAATAAAACAATAGTCTTT
TTTGTCCATCAATGTTTTAATTCTCTTGTTAAACTGTTATTTTGTTATCAGAAA

ACGTCATCCTCAGGAGGGGACCCAGAAATTGTACTTCACAGTTTTAATTGTGGA
TGCAGTAGGAGTCCTCCCCTGGGTCTTTAACATGAAGTGTCAAAATTAACACCT

GGGGAATTTTTCTACTGTAATACAACACAACTGTTTAATAGTACTTGGAATAGT
CCCCTTAAAAAGATGACATTATGTTGTGTTGACAAATTATCATGAACCTTATCA

ACTGAAGGGTCAAATAACACTGGAGGAAATGACACAATCACACTCCCATGCAGA
TGACTTCCCAGTTTATTGTGACCTCCTTTACTGTGTTAGTGTGAGGGTACGTCT

ATAAAACAAATTGTAAACATGTGGCAGGAAGTAGGAAAAGCAATGTATGCCCCT
TATTTTGTTTAACATTTGTACACCGTCCTTCATCCTTTTCGTTACATACGGGGA

CCCATCAGTGGACAAATTAAATGTATATCAAATATTACAGGGCTACTATTAACA
GGGTAGTCACCTGTTTAATTTACATATAGTTTATAATGTCCCGATGATAATTGT

AGAGATGGGGGTGAAGATACAACTAATACTACAGA
TCTCTACCCCCACTTCTATGTTGATTATGATGTCTCTAG

TABLE 4B

MetLeuArgProValGluThrProThrArgGluIleLysLysLeuAspGlyLeu
TrpAlaPheSerLeuAspArgGluArgValAlaAspLeuAsnAlaSerValGln
IleAsnCysThrArgProAsnAsnAsnThrArgLysSerIleThrLysGlyPro
GlyArgValIleTyrAlaThrGlyGlnIleIleGlyAspIleArgLysAlaHis
CysAsnLeuSerArgAlaGlnTrpAsnAsnThrLeuLysGlnValValThrLys
LeuArgGluGlnPheAspAsnLysThrIleValPheThrSerSerSerGlyGly
AspProGluIleValLeuHisSerPheAsnCysGlyGlyGluPhePheTyrCys
AsnThrThrGlnLeuPheAsnSerThrTrpAsnSerThrGluGlySerAsnAsn
ThrGlyGlyAsnAspThrIleThrLeuProCysArgIleLysGlnIleValAsn
MetTrpGlnGluValGlyLysAlaMetTyrAlaProProIleSerGlyGlnIle
LysCysIleSerAsnIleThrGlyLeuLeuLeuThrArgAspGlyGlyGluAsp
ThrThrAsnThrThrGluIleArgArgGlnAlaSerArgGluLeuGluPheLeu
LysThrLysGlyProArgAspThrProIlePheIleGly

TABLE 4C

ATGTTACGTCCTGTAGAAACCCCAACCCGTGAAATCAAAAAACTGGACGGCCTG

TGGGCATTCAGTCTGGATCGCGAACGCGTGGCCGATCTGAATGCATCTGTACAA

ATTAATTGTACAAGACCCAACAACAATACAAGAAAAAGTATAACTAAGGGACCA

GGGAGAGTAATTTATGCAACAGGACAAATAATAGGAGATATAAGAAAAGCACAT

TGTAACCTTAGTAGAGCACAATGGAATAACACTTTAAAACAGGTAGTTACAAAA

TTAAGAGAACAATTTGACAATAAAACAATAGTCTTTACGTCATCCTCAGGAGGG

GACCCAGAAATTGTACTTCACAGTTTTAATTGTGGAGGGGAATTTTTCTACTGT

AATACAACACAACTGTTTAATAGTACTTGGAATAGTACTGAAGGGTCAAATAAC

ACTGGAGGAAATGACACAATCACACTCCCATGCAGAATAAAACAAATTGTAAAC

ATGTGGCAGGAAGTAGGAAAAGCAATGTATGCCCCTCCCATCAGTGGACAAATT

AAATGTATATCAAATATTACAGGGCTACTATTAACAAGAGATGGGGGTGAAGAT

ACAACTAATACTACAGAGATCCGTCGACAAGCTTCCCGGGAGCTGGAATTCTTG

AAGACGAAAGGGCCTCGTGATACTCCTATTTTTATAGGT

TABLE 5

GluAsnPheThrAspAsnAlaLysThrIleIleValHisLeuAsnGlu
SerValGlnIleAsnCysThrArgProAsnTyrAsnLysArgLysArgIleHis
IleGlyProGlyArgAlaPheTyrThrThrLysAsnIleIleGlyThrIleArg
GlnAlaHisCysAsnIleSerArgAlaLysTrpAsnAspThrLeuArgGlnIle
ValSerLysLeuLysGluGlnPheLysAsnLysThrIleValPheAsnGlnSer
SerGlyGlyAspProGluIleValMetHisSerPheAsnCysGlyGlyGluPhe
PheTyrCysAsnThrSerProLeuPheAsnSerThrCysLysIleLysGlnIle
IleAsnMetTrpGlnGluValGlyLysAlaMetTyrAlaProProIleGluGly
GlnIleArgCysSerSerAsnIleThrGlyLeuLeuLeuThrArgAspGlyGly
LysAspThrAspThrAsnAspThr

TABLE 5A

```
GATCTGAAAATTTCACAGACAATGCTAAAACCATAATAGTACACCTGAATGAA
    ACTTTTAAAGTGTCTGTTACGATTTTGGTATTATCATGTGGACTTACTT

TCTGTACAAATTAATTGTACAAGACCCAACTACAATAAAAGAAAAAGGATACAT
AGACATGTTTAATTAACATGTTCTGGGTTGATGTTATTTTCTTTTTCCTATGTA

ATAGGACCAGGGAGAGCATTTTATACAACAAAAAATATAATAGGAACTATAAGA
TATCCTGGTCCCTCTCGTAAAATATGTTGTTTTTTATATTATCCTTGATATTCT

CAAGCACATTGTAACATTAGTAGAGCAAAATGGAATGACACTTTAAGACAGATA
GTTCGTGTAACATTGTAATCATCTCGTTTTACCTTACTGTGAAATTCTGTCTAT

GTTAGCAAATTAAAAGAACAATTTAAGAATAAAACAATAGTCTTTAATCAATCC
CAATCGTTTAATTTTCTTGTTAAATTCTTATTTTGTTATCAGAAATTAGTTAGG

TCAGGAGGGGACCCAGAAATTGTAATGCACAGTTTTAATTGTGGAGGGGAATTT
AGTCCTCCCCTGGGTCTTTAACATTACGTGTCAAAATTAACACCTCCCCTTAAA

TTCTACTGTAATACATCACCACTGTTTAATAGTACTTGCAAAATAAAACAAATT
AAGATGACATTATGTAGTGGTGACAAATTATCATGAACGTTTTATTTTGTTTAA

ATAAACATGTGGCAGGAAGTAGGAAAAGCAATGTATGCCCCTCCCATTGAAGGA
TATTTGTACACCGTCCTTCATCCTTTTCGTTACATACGGGGAGGGTAACTTCCT

CAAATTAGATGTTCATCAAATATTACAGGGCTACTATTAACAAGAGATGGTGGT
GTTTAATCTACAAGTAGTTTATAATGTCCCGATGATAATTGTTCTCTACCACCA

AAGGACACGGACACGAACGACACCGA
TTCCTGTGCCTGTGCTTGCTGTGGCTCTAG
```

TABLE 5B

MetLeuArgProValGluThrProThrArgGluIleLysLysLeuAspGlyLeu
TrpAlaPheSerLeuAspArgGluArgValAlaAspIleIleAspGlySerGlu
AsnPheThrAspAsnAlaLysThrIleIleValHisLeuAsnGluSerValGln
IleAsnCysThrArgProAsnTyrAsnLysArgLysArgIleHisIleGlyPro
GlyArgAlaPheTyrThrThrLysAsnIleIleGlyThrIleArgGlnAlaHis
CysAsnIleSerArgAlaLysTrpAsnAspThrLeuArgGlnIleValSerLys
LeuLysGluGlnPheLysAsnLysThrIleValPheAsnGlnSerSerGlyGly
AspProGluIleValMetHisSerPheAsnCysGlyGlyGluPhePheTyrCys
AsnThrSerProLeuPheAsnSerThrCysLysIleLysGlnIleIleAsnMet
TrpGlnGluValGlyLysAlaMetTyrAlaProProIleGluGlyGlnIleArg
CysSerSerAsnIleThrGlyLeuLeuLeuThrArgAspGlyGlyLysAspThr
AspThrAsnAspThrGluIleArgArgGlnAlaSerArgGluLeuGluPheLeu
LysThrLysGlyProArgAspThrProIlePheIleGly

TABLE 5C

ATGTTACGTCCTGTAGAAACCCCAACCCGTGAAATCAAAAAACTGGACGGCCTG

TGGGCATTCAGTCTGGATCGCGAACGCGTGGCCGATATCATCGATGGATCTGAA

AATTTCACAGACAATGCTAAAACCATAATAGTACACCTGAATGAATCTGTACAA

ATTAATTGTACAAGACCCAACTACAATAAAAGAAAAAGGATACATATAGGACCA

GGGAGAGCATTTTATACAACAAAAAATATAATAGGAACTATAAGACAAGCACAT

TGTAACATTAGTAGAGCAAAATGGAATGACACTTTAAGACAGATAGTTAGCAAA

TTAAAAGAACAATTTAAGAATAAAACAATAGTCTTTAATCAATCCTCAGGAGGG

GACCCAGAAATTGTAATGCACAGTTTTAATTGTGGAGGGGAATTTTTCTACTGT

AATACATCACCACTGTTTAATAGTACTTGCAAAATAAAACAAATTATAAACATG

TGGCAGGAAGTAGGAAAAGCAATGTATGCCCCTCCCATTGAAGGACAAATTAGA

TGTTCATCAAATATTACAGGGCTACTATTAACAAGAGATGGTGGTAAGGACACG

GACACGAACGACACCGAGATCCGTCGACAAGCTTCCCGGGAGCTGGAATTCTTG

AAGACGAAAGGGCCTCGTGATACTCCTATTTTTATAGGT

TABLE 6

LeuLysGluAlaValGluIleAsnCysThrArgProAsnAsnAsnThrThrArg

SerIleHisIleGlyProGlyArgAlaPheTyrAlaThrGlyAspIleIleGly

AspIleArgGlnAlaHisCysAsnIleSerArgAlaLysTrpAsnAsnThrLeu

LysGlnIleValIleLysLeuArgAspGlnPheGluAsnLysThrIleIlePhe

AsnArgSerSerGlyGlyAspProGluIleValMetHisSerPheAsnCysGly

GlyGluPhePheTyrCysAsnSerThrGlnLeuPheSerSerThrTrpAsnGly

ThrGluGlySerAsnAsnThrGlyGlyAsnAspThrIleThrLeuProCysArg

IleLysGluIleIleAsnMetTrpGlnGluValGlyLysAlaMetTyrAlaPro

ProIleLysGlyGlnValLysCysSerSerAsnIleThrGlyLeuLeuLeuThr

ArgAspGlyGlyAsnSerLysAsnGlySerLysAsnThr

TABLE 6A

CTGAAAGAAGCTGTAGAAATTAATTGTACAAGGCCCAACAACAATACAACAAGA
GACTTTCTTCGACATCTTTAATTAACATGTTCCGGGTTGTTGTTATGTTGTTCT

AGTATACATATAGGACCAGGGAGAGCATTTTATGCAACAGGAGACATAATAGGA
TCATATGTATATCCTGGTCCCTCTCGTAAAATACGTTGTCCTCTGTATTATCCT

GATATAAGACAAGCACATTGTAACATTAGTAGAGCAAAATGGAATAACACTTTA
CTATATTCTGTTCGTGTAACATTGTAATCATCTCGTTTTACCTTATTGTGAAAT

AAACAGATAGTTATAAAATTAAGAGACCAATTTGAGAATAAAACAATAATCTTT
TTTGTCTATCAATATTTTAATTCTCTGGTTAAACTCTTATTTTGTTATTAGAAA

AATCGATCCTCAGGAGGAGACCCAGAAATTGTAATGCACAGTTTTAATTGTGGA
TTAGCTAGGAGTCCTCCTCTGGGTCTTTAACATTACGTGTCAAAATTAACACCT

GGGGAATTTTTCTACTGTAATTCAACACAACTGTTTAGTAGTACTTGGAATGGT
CCCCTTAAAAAGATGACATTAAGTTGTGTTGACAAATCATCATGAACCTTACCA

ACTGAAGGGTCAAATAACACTGGAGGAAATGACACAATCACCCTCCCATGCAGA
TGACTTCCCAGTTTATTGTGACCTCCTTTACTGTGTTAGTGGGAGGGTACGTCT

ATAAAAGAAATTATAAACATGTGGCAGGAAGTAGGAAAAGCAATGTATGCCCCT
TATTTTCTTTAATATTTGTACACCGTCCTTCATCCTTTTCGTTACATACGGGGA

CCCATCAAAGGACAAGTTAAATGTTCATCAAATATTACAGGGCTGCTATTAACA
GGGTAGTTTCCTGTTCAATTTACAAGTAGTTTATAATGTCCCGACGATAATTGT

AGAGATGGTGGTAATAGCAAGAATGGTAGCAAGAATACAGA
TCTCTACCACCATTATCGTTCTTACCATCGTTCTTATGTCTCTAG

TABLE 6B

MetLeuArgProValGluThrProThrArgGluIleLysLysLeuAspGlyLeu

TrpAlaPheSerLeuAspArgGluArgValAlaAspLeuLysGluAlaValGlu

IleAsnCysThrArgProAsnAsnAsnThrThrArgSerIleHisIleGlyPro

GlyArgAlaPheTyrAlaThrGlyAspIleIleGlyAspIleArgGlnAlaHis

CysAsnIleSerArgAlaLysTrpAsnAsnThrLeuLysGlnIleValIleLys

LeuArgAspGlnPheGluAsnLysThrIleIlePheAsnArgSerSerGlyGly

AspProGluIleValMetHisSerPheAsnCysGlyGlyGluPhePheTyrCys

AsnSerThrGlnLeuPheSerSerThrTrpAsnGlyThrGluGlySerAsnAsn

ThrGlyGlyAsnAspThrIleThrLeuProCysArgIleLysGluIleIleAsn

MetTrpGlnGlyValGlyLysAlaMetTyrAlaProProIleLysGlyGlnVal

LysCysSerSerAsnIleThrGlyLeuLeuLeuThrArgAspGlyGlyAsnSer

LysAsnGlySerLysAsnThrGluIleArgArgGlnAlaSerArgGluLeuGlu

PheLeuLysThrLysGlyProArgAspThrProIlePheIleGly

TABLE 6C

ATGTTACGTCCTGTAGAAACCCCAACCCGTGAAATCAAAAAACTGGACGGCCTG
TGGGCATTCAGTCTGGATCGCGAACGCGTGGCCGATCTGAAAGAAGCTGTAGAA
ATTAATTGTACAAGGCCCAACAACAATACAACAAGAAGTATACATATAGGACCA
GGGAGAGCATTTTATGCAACAGGAGACATAATAGGAGATATAAGACAAGCACAT
TGTAACATTAGTAGAGCAAAATGGAATAACACTTTAAAACAGATAGTTATAAAA
TTAAGAGACCAATTTGAGAATAAAACAATAATCTTTAATCGATCCTCAGGAGGA
GACCCAGAAATTGTAATGCACAGTTTTAATTGTGGAGGGGAATTTTTCTACTGT
AATTCAACACAACTGTTTAGTAGTACTTGGAATGGTACTGAAGGGTCAAATAAC
ACTGGAGGAAATGACACAATCACCCTCCCATGCAGAATAAAAGAAATTATAAAC
ATGTGGCAGGAAGTAGGAAAAGCAATGTATGCCCCTCCCATCAAAGGACAAGTT
AAATGTTCATCAAATATTACAGGGCTGCTATTAACAAGAGATGGTGGTAATAGC
AAGAATGGTAGCAAGAATACAGAGATCCGTCGACAAGCTTCCCGGGAGCTGGAA
TTCTTGAAGACGAAAGGGCCTCGTGATACTCCTATTTTTATAGGT

TABLE 7

LeuAsnGluSerValGluIleAsnCysThrArgProTyrAsnAsnValArgArg

SerLeuSerIleGlyProGlyArgAlaPheArgThrArgGluIleIleGlyIle

IleArgGlnAlaHisCysAsnIleSerArgAlaLysTrpAsnAsnThrLeuLys

GlnIleValGluLysLeuArgGluGlnPheLysAsnLysThrIleValPheAsn

HisSerSerGlyGlyAspProGluIleValThrHisSerPheAsnCysGlyGly

GluPhePheTyrCysAsnSerThrGlnLeuPheAsnSerThrTrpAsnGlyThr

AspIleLysGlyAspAsnLysAsnSerThrLeuIleThrLeuProCysArgIle

LysGlnIleIleAsnMetTrpGlnGlyValGlyLysAlaMetTyrAlaProPro

IleGlnGlyGlnIleArgCysSerSerAsnIleThrGlyLeuLeuLeuThrArg

AspGlyGlyAsnSerSerSerArgGlu

TABLE 7A

CTGAATGAATCTGTAGAAATTAATTGTACAAGACCCTACAACAATGTAAGAAGA
GACTTACTTAGACATCTTTAATTAACATGTTCTGGGATGTTGTTACATTCTTCT

AGTCTATCTATAGGACCAGGGAGAGCATTTCGTACAAGAGAAATAATAGGAATT
TCAGATAGATATCCTGGTCCCTCTCGTAAAGCATGTTCTCTTTATTATCCTTAA

ATAAGACAAGCACATTGTAACATTAGTAGAGCAAAATGGAATAACACTTTAAAA
TATTCTGTTCGTGTAACATTGTAATCATCTCGTTTTACCTTATTGTGAAATTTT

CAGATAGTTGAGAAATTAAGAGAACAATTTAAGAATAAAACAATAGTCTTTAAT
GTCTATCAACTCTTTAATTCTCTTGTTAAATTCTTATTTTGTTATCAGAAATTA

CATTCCTCAGGAGGGGACCCAGAAATTGTAACGCACAGTTTTAATTGTGGAGGG
GTAAGGAGTCCTCCCCTGGGTCTTTAACATTGCGTGTCAAAATTAACACCTCCC

GAATTTTTCTACTGTAATTCAACACAACTGTTTAATAGTACTTGGAATGGTACT
CTTAAAAAGATGACATTAAGTTGTGTTGACAAATTATCATGAACCTTACCATGA

GACATTAAAGGAGATAATAAAAATAGCACACTCATCACACTCCCATGCAGAATA
CTGTAATTTCCTCTATTATTTTTATCGTGTGAGTAGTGTGAGGGTACGTCTTAT

AAACAAATTATAAACATGTGGCAGGGAGTAGGCAAAGCAATGTATGCCCCTCCC
TTTGTTTAATATTTGTACACCGTCCCTCATCCGTTTCGTTACATACGGGGAGGG

ATCCAAGGACAAATTAGATGTTCATCAAATATTACAGGGCTGCTATTAACAAGA
TAGGTTCCTGTTTAATCTACAAGTAGTTTATAATGTCCCGACGATAATTGTTCT

GATGGTGGTAATAGCAGCAGCAGGGAAGA
CTACCACCATTATCGTCGTCGTCCCTTCTCTAG

TABLE 7B

MetLeuArgProValGluThrProThrArgGluIleLysLysLeuAspGlyLeu

TrpAlaPheSerLeuAspArgGluArgValAlaAspLeuAsnGluSerValGlu

IleAsnCysThrArgProTyrAsnAsnValArgArgSerLeuSerIleGlyPro

GlyArgAlaPheArgThrArgGluIleIleGlyIleIleArgGlnAlaHisCys

AsnIleSerArgAlaLysTrpAsnAsnThrLeuLysGlnIleValGluLysLeu

ArgGluGlnPheLysAsnLysThrIleValPheAsnHisSerSerGlyGlyAsp

ProGluIleValThrHisSerPheAsnCysGlyGlyGluPhePheTyrCysAsn

SerThrGlnLeuPheAsnSerThrTrpAsnGlyThrAspIleLysGlyAspAsn

LysAsnSerThrLeuIleThrLeuProCysArgIleLysGlnIleIleAsnMet

TrpGlnGlyValGlyLysAlaMetTyrAlaProProIleGlnGlyGlnIleArg

CysSerSerAsnIleThrGlyLeuLeuLeuThrArgAspGlyGlyAsnSerSer

SerArgGluGluIleArgArgGlnAlaSerArgGluLeuGluPheLeuLysThr

LysGlyProArgAspThrProIlePheIleGly

TABLE 7C

ATGTTACGTCCTGTAGAAACCCCAACCCGTGAAATCAAAAAACTGGACGGCCTG
TGGGCATTCAGTCTGGATCGCGAACGCGTGGCCGATCTGAATGAATCTGTAGAA
ATTAATTGTACAAGACCCTACAACAATGTAAGAAGAAGTCTATCTATAGGACCA
GGGAGAGCATTTCGTACAAGAGAAATAATAGGAATTATAAGACAAGCACATTGT
AACATTAGTAGAGCAAAATGGAATAACACTTTAAAACAGATAGTTGAGAAATTA
AGAGAACAATTTAAGAATAAAACAATAGTCTTTAATCATTCCTCAGGAGGGGAC
CCAGAAATTGTAACGCACAGTTTTAATTGTGGAGGGGAATTTTTCTACTGTAAT
TCAACACAACTGTTTAATAGTACTTGGAATGGTACTGACATTAAAGGAGATAAT
AAAAATAGCACACTCATCACACTCCCATGCAGAATAAAACAAATTATAAACATG
TGGCAGGGAGTAGGCAAAGCAATGTATGCCCCTCCCATCCAAGGACAAATTAGA
TGTTCATCAAATATTACAGGGCTGCTATTAACAAGAGATGGTGGTAATAGCAGC
AGCAGGGAAGAGATCCGTCGACAAGCTTCCCGGGAGCTGGAATTCTTGAAGACG
AAAGGGCCTCGTGATACTCCTATTTTTATAGGT

Table 8

```
             5'   CTGAACCAATCTGTAGAAATTAATTGTACAAGACCCAAC
                  GACTTGGTTAGACATCTTTAATTAACATGTTCTGGGTTG

AACAATACAAGAAAAAGTATCCGTATCCAGAGAGGACCAGGGAGAGCATTTGTT
TTGTTATGTTCTTTTTCATAGGCATAGGTCTCTCCTGGTCCCTCTCGTAAACAA

ACAATAGGAAAAATAGGAAATATGAGACAAGCACATTGTAACATTAGTAGAGCA
TGTTATCCTTTTTATCCTTTATACTCTGTTCGTGTAACATTGTAATCATCTCGT

AAATGGAATAACACTTTAAAACAGATAGATAGCAAATTAAGAGAACAATTTGGA
TTTACCTTATTGTGAAATTTTGTCTATCTATCGTTTAATTCTCTTGTTAAACCT

AATAATAAAACAATAATCTTTAAGCAGTCCTCAGGAGGGGACCCAGAAATTGTA
TTATTATTTTGTTATTAGAAATTCGTCAGGAGTCCTCCCCTGGGTCTTTAACAT

ACGCACAGTTTTAATTGTGGAGGGGAATTTTTCTACTGTAATTCAACACAACTG
TGCGTGTCAAAATTAACACCTCCCCTTAAAAAGATGACATTAAGTTGTGTTGAC

TTTAATAGTACTTGGTTTAATAGTACTTGGAGTACTAAAGGGTCAAATAACACT
AAATTATCATGAACCAAATTATCATGAACCTCATGATTTCCCAGTTTATTGTGA

GAAGGAAGTGACACAATCACCCTCCCATGCAGAATAAAACAAATTATAAACATG
CTTCCTTCACTGTGTTAGTGGGAGGGTACGTCTTATTTTGTTTAATATTTGTAC

TGGCAGGAAGTAGGAAAAGCAATGTATGCCCCTCCCATCAGTGGACAAATTAGA
ACCGTCCTTCATCCTTTTCGTTACATACGGGGAGGGTAGTCACCTGTTTAATCT

TGTTCATCAAATATTACAGGGCTGCTATTAACAAGAGATGGTGGTAATAGCAAC
ACAAGTAGTTTATAATGTCCCGACGATAATTGTTCTCTACCACCATTATCGTTG

AATGAGTCCGA          3'
TTACTCAGGCTCTAG
```

Table 9.

EE1-707-1	CTRPNNY TRK RIHI	GPGRAFY TTGQVIGRI RQAQC
MN	CTRPNYN KRK RIHI	GPGRAFY TTKNIIGTI RQAHC
EE1-708-1	CTRPNNN TRR RIHI	GPGRAFY TTGQVIGRI RQAQC
KW4-4-2	CTRPNNN TRR RIHI	GPGRAFY TMGQIIGDI RKAHC
JH3	CTRPSKT TRR RIHI	GPGRAFY ATGDIAGDL RQAHC
WMJ3	CTRPNDI ARR RIHI	GPGRAFY T GKIIGNI RQAHC
BAL	CTRPNNN TRK SIHI	GPGRAFY TTGEIIGDI RQAHC
EE3-7-3	CTIPNNN TRK SIHI	GPGRAFY TTGEIIGDI RQAHC
KW4-6-1	CTRPNNN TRK SIHI	GPGRAFY TTGEIIGDI RQAHC
EE3-4-3	CTRPNNN TRK SIHI	GPGRAFY TTGEIIGDI RQAHC
EE7-3-1	CTRPNNN TRK SIHI	GPGRAFY TTGEIIGNI RQAHC
EE7-15-3	CTRPNNN TRK SIHI	GPGRAFY TTGEIIGDI RQAYC
EE7-24-1	CTRPNNN TRK SIHI	GPGRAFY TTGEIIGDI RQAHC
EE6-4-1	CTRPNNN TRK SIHI	GPGRAFY ATGAIIGDI RQAHC
JG1	CTRPNNN TRK SIHI	GPGRAFY ATGDIIGDI RQAHC
DD3-1	CTRPNNN TRK GIHI	GPGRAFY ATGDIIGDI RQAHC
DD7-1	CTRPNNN TRK GIHI	GPGRAFY ATGDIIGDI RQAHC
EE1-330-1	CIRPNNN TRK GIHI	GPGRAFY ATGDIIGDI RQAHC
SC	CTRPNNN TTR SIHI	GPGRAFY ATGDIIGDI RQAHC
DD10-1	CTRPNNNVRRR HIHI	GPGRAFY TGEIRGNI RQAHC
WMJ1.5	CTRPNNNVRRR HIHI	GPGRAFYY GEIRGNI RQAHC
KW3-2-2	CIRPHNTI RR RIHI	GPGRAFS TTRGIQGDI RQAYC
DD4-1	CTRPYSNV RN RIHI	GPGRAFH TTKRITGDM RQARC
AFL30-4-1	CTRPYSNV RN RIHI	GPGRAFH TTKRITGDM RQARC
KW2-1-1	CTRPYSNV RN RIHI	GPGRAFH TTKRITGDM RQARC
KW2-9-1	CTRPYSNV RN RIHI	GPGRAFH TTKRITGDM RQARC
EE6-4-4	CTRPYSNV RN RIHI	GPGRAFH TTKRITGDM RQARC
EE7-20-3	XTRPYSNV RN RIHI	GPGRAFH TTKRITGDM RQARC
RJS426	CTRPYNKK RIRHMHI	GPGRAFY ATG GMGDI RQAHC
KW4-13-1	CTRPNNKIPR HFHI	GPGRAFY ATGGIEGDI RKARC
AFL30-6-2	XXRPNNY TRK GIHI	GPGRAIY ATGDIIGDI RQAHC
KW2-8-2	XTRPNNN TRK GIHI	GPGRAVY TTGRIVGDI RQAHC
TM5-16-1	CTRPNNN TRK GIHI	GPGRAVY TTGRIVGDI RQAHC
KW2-8-1	CTRPDNN ARK GIHI	GPGRAVY TTGRIIGDI RQAHC
EE5-6-1	CTRPDNN ARK GIHI	GPGRAVY TTGRIIGDI RQAHC
EE6-3-5	XTRPNNN TKK GIHI	GPGRAWY TRTRIIGDI RQAHC
TM5-8-1	CTRPNNN TKK GIHI	GPGRÂVY TARRIIGDI RQAHC
TM5-12-1	CIRPNIN TGK RIPI	GPGRAFY TTGAIKGNI RQAHC
EE5-3-2	XTRPNNN TRK SIPI	GPGRAFY TTGEIIGDI RQAHC
EE3-2-4	XXXPSNN TRK SIPI	GPGRAFY ATGDIIGNI RQAHC
WH331	XXXXXXX XRK SIPI	GPGRAFY ATGEIIGDI RQAHC
EE1-279-1	CTRPNNN TRK RISI	GPGRAFY ATRQIVGDI RQAHC
EE3-5-1	CMRPNNN TRK SINI	GPGRAFY TTGQIIGDI RQAHC
WH721	XXXXNNN TRK SINI	GPGRAFY ARGEIIGDI RRAXX
KW4-3-2	CTRPNNN TRK SIAI	GPGRAFY ATRRIIGDI RQAHC
TM4-11-1	CTRPNNN TRK RIYI	GPGRAFY ARQQIIGDI RQAHC
TM5-13-1	CTRPNNN TRK RIYI	GPGRAFY ARQQVIGDI RQAHC
SF2	CTRPNNN TRK SIYI	GPGRAFH TTGRIIGDI RKAHC
SF4	CTRPNNN TRK SIYI	GPGRAFH TTGRIIGDI RKAHC
Z321	CMRPNNN TRK SISI	GPGRAFF ATGDIIGDI RQAHC
WMJ2	CTRPYNNV RR SLSI	GPGRAFR YRE IIGII RQAHC
EE6-1-1	CTRPNNN TRK RIKI	GPGRAFV TTRQIIGDI RQAYC
KW4-10-1	CTRPNNN TKK GIYI	GPGRAVY TTEKIIGDI RRAHC
TM5-7-1	CTRPNNN TKK GIYI	GPGRAVY TTEKIIGDI RRAHC
EE7-15-1	CTRPNNN TKK GIRI	GPGRAVY TAQRIIGDI RQAHC

Table 9 (continued)

DD-3-3	CTRPNNN TKK GIRI GPGRAVY TARRIIGDI RQAHC
AFL30-5-3	CTRPNNN TRK GIRI GPGRAIY ATARIIGDI RQAHC
KW3-9-2	CIRPNNN TRK RIGI GPGRAIL QQ ENIGDI RQAHC
JG2	CTRPNNN TRK SINI GPGRAFY ATGQIIGNI RQAHC
EE3-5-2	CHRPNNN TRK SINI GPGRALY TTGQIIGDI RQAHC
AFL30-1-3	XTRPNNN TSR GIRI GPGRAIL ATERIIGDI RQAHC
AFL30-3-1	CTRPNNN TSR GIRI GPGRAIL ATERIIGDI RQAHC
KW2-2-2	CTRPNNN TSR GIRI GPGRAIL ATERIIGDI RQAHC
EE5-3-3	XTRPNNN TSR GIRI GPGRAIL ATERIIGDI RQAHC
EE5-6-3	CTRPNNN TSR GIRI GPGRAIL ATERIIGDI RQAHC
EE5-10-3	CTRPNNN TSR GIRI GPGRAIL ATERIIGDI RQAHC
EE5-11-3	CTRPNNN TSR GIRI GPGRAIL ATERIIGDI RQAHC
EE7-24-3	CTRPNNN TSR GIRI GPGRAIL ATERIIGDI RQAHC
KW4-3-1	CTRPNNN TKR GIRI GPGRAVY ATDRIIGDI RQAHC
KW4-2-2	CTRPNNN TKR GIRI GPGRAVM QQTRIIGDI RQAHC
WH425	XXXXNNN TRR SISI GPGRALY TTGAIIGSI RQAXX
WH718	XXXXXXX XSR SISI GPGRAWY QQEEVIGDI XXXXX
WH244	XXXXNNKIKIR RIHI GPGRPFY TTK IGDI RQAYC
EE3-6-2	CTRPNNN TRK GIHI GPGRTFY ATGAIIGDI RQAHC
EE6-3-1	CTRPNNN TKK GIHI GPGRKWY TRTKIIGDI RQAHC
DD9-1	CTRPNNN TRR SIHI GPGRWSVHTTGEIVGDI RQAHC
TM5-11-1	CTRPNNN TQK RITI GPGRVFY TTGKIVGDI RQAHC
EE5-10-5	CTRPNNN TRK GIFI GPGRNIY TTGNIIGDI RKAHC
EE5-11-1	CTRPNNN TRK GIFI GPGRNIY TTGNIIGDI RKAHC
EE5-14-1	CTRPNNN TRK GIFI GPGRNIY TTGNIIGDI RKAHC
KW2-6-1	CTRPNNN TRK GIFI GPGRNIY TTGNIIGDI RKAHC
DD5-1	XXXXNNN TRA RLSI GPGRSFY ATRNIVGDI RQAHC
TM4-7-3	CIRPNNN TRK AMSI GPGRKLY TRNKIIGDI RQAHC
NY5	CTRPNNN TKK GIAI GPGRTLY AREKIIGDI RQAHC
AFL30-11-1	CTRPNEN TKK SLYI GPGRRFH VTKAITGDI RQAHC
AFL30-12-1	CTRPNEN TKR SLYI GPGRRFH VTKAITGDI RQAHC
KW4-12-1	CTRPSNN TKQ SLYI GPGRRFH VTKAITGDI RQAHC
KW4-12-2	CTRPSNN TKR SLYI GPGRRFH VTKAITGDI RQAHC
KW4-7-1	CTRPNNN TRK GIHM GPGRAFY TQENI GDI RQARC
KW4-7-2	CTRPNNN TRK GIHM GPGRAFY TQENI GDI RQAHC
KW3-6-1	CIRPNNI TRR SMSM GPGRAFV ATRQIIGDI RKAHC
IIIB (BH10)	CTRPNNN TRK SIRIQRGPGRAFV TIGK IGNM RQAHC
LAV-BRU	CTRPNNN TRK SIHIQRGPGRAFV TIGK IGNM RQAHC
EE5-10-1	CTRPNNN TRK SIRIQRGPGRAFV TIGK IXNM RQAHC
EE7-3-3	CTRPNNN TRK SIRIQRGPGRAFV TIGK IGNM RQAHC
EE7-6-2	CTRPNNN TRK SIRIQRGPGRAFV TIGK IGNM RQAHC
KW2-1-3	CTRPNNN TRK RIRIQRGPGRAFV TIGK IGNMEEQAHC
EE3-6-1	CTRPNNN TRK RIRIQRGPGRAFV TIGK IGNM RQAHC
EE7-15-2	CTRPNNN TRK KIRIQRGPGRAFV TIGK IRNM RQAHC
EE5-3-1	CTRPNNN TRK KIRIQRGPGRAFV TIGK IRNM RQAHC
TM4-14-2	CTRPNNN TRK KIVSRGGPGRAFV TIGK IRNM RQAHC
KW3-6-2	CTRPNNN TRR GIRV GPGRAVY STDKIIGDI RQAHC
RF/TM3-8-1	CTRPNNN TRK RITK GPGRVIY ATGQIIGDI RKAHC
RF	CTRPNNN TRK SITK GPGRVIY ATGQIIGDI RKAHC
EE7-20-1	CTRPNNH TEK RITL GPGRVLY TTGRIIGDI RRAHC
CDC451	CTRPNNH TRK RVTL GPGRVWY TTGEILGNI RQAHC
TM3-9-4	CTRPNNN TRK RVTM GPGRVWY TTGEIVGDI RQAHC
EE1-706-1	CTRPNNN TRK RITM GPGRVYY TTGQIIGNI RQAHC
EE1-706-2	CTRPNNN TRK RITM GPGRVYY TAGQIIGNI RQAHC
BRVA	CTRPNNN TRK RITM GPGRVYY TTGQIIGDI RRAHC

Table 9 (continued)

KW2-1-2	CTRPNNN TRK RITM GPGRVYY TTGQIIGNI RQAHC
EE1-317-1	CTRPNNN TRK GIHI GPGT FY TTGEIIGDI RQAHC
SF170	CTRPNNN TRK SGTI GPGQAFY ATGDIIGDI RQAYC
KW2-6-2	CTKPNNN TKK GIFI GPGKNIY TTENIIGDI RKAHC
Z3	CTRPGSDKKIRQSIRI GPGKVFY AKGGITG QAHC
EE3-6-1	CIRPNNN TRK SIPM GPGKAFY ATGEIIGDI RQAHC
KW4-8-1	CTRPNNN TRK SIPM GPGKAFY TTGDIIGDI RQAHC
KW4-8-3	CTRPNNN TRK SIPM GPGKAFY ATGDIIGDI RQAHC
TM4-14-1	CTRPNNY ARR GIRV GPGSATY TAPSIIADI SQAHC
KW2-2-1	CTRPNNN TEK ESYQRGPGGAFV TIGK IGNM RQAHC
LAV-MAL	CTRPGNN TRR GIHF GPGQALY TTGIVIGDI RRAYC
EE1-317-2	CTRPNNN IRK RITR GPGKVIY ATGQIIGDI RKAHC
EE1-707-2	CTRPNNN IRK RITR GPGKVIY ATGQIIGDI RKAHC
SF33	CTRPNNN RRR RITS GPGKVLY TTGEIIGDI RKAYC
EE7-21-3	CTRPNNN TRR RIHI GPRRAFY TTGQVIGRI RQAQC
TM4-14-3	CTRPNNY AKR GIRV EPGKAII ATKKIIENI KKAHY
~~EE~~7-6-1	CIRPGNN TRR SISI RPERAFFTQTGDVIGDI RQAHC
DD1-1	CTRPNNN TRK SIRI GAGRAIY ATARIIGDI RQAHC
AFL30-5-1	CTRPNNN TRK SIRI GAGRAIY ATERIIGDI RQAHC
EE7-21-1	CTRPNNN TRK SIRY GAGRAIY ATARIIGDI RQAHC
TM4-13-2	CIRPNNN TKR AIYI GQGRAIH TTDRIIGDI RQAHC
TM4-13-3	CTRPNNN TKR AIYI GQGRAIH TTDRIXGDI RQAHC
TM3-7-1	XIRPNNN TRK GIYV GSGRAVY TRDKIMGDI RQAHC
AFL30-9-1	CIRPNNN TRK GIYV GSGRKVY TRHKIIGDI RQAHC
AFL30-9-2	CIRPNNN TRK GIYV GSGRKVY TRQKIIGDI RQAHC
LAVELI	CARPYQN TRQ RTPI GLGQSLY TTRSRS II GQAHC
JY1	CTRPDNKITRQ STPI GLGQALY TTR IKGDI RQAYC
Z6	CTRPYKN TRQ STPI GLGQALY TTRGRTKII GQAHC

Table 10

x_{13}	x_{12}	x_{11}	x_{10}	x_9	x_8	x_7	x_6	x_5	x_4	x_3	x_2	x_1			
T 115	R 129	P 132	N115	N 121	N 118	T 113	R 103	K 90	R 37	I 120	H 47	I 100	G 136	P 127	G 136
I 11	I 1		Y 10	S 7	NV 2	K 3	K 18	R 34	S 47	L 6	R 28	R 2	E 1	L 3	E 1
M 3	K 1		S 4	K 2	Y 4	V 7	S 9	I 1	G 43	M 3	Y 15	M 11	R 1	A 3	R 1
A 1			D 3	E 2	KI 3	I 3	I 2	N 6	H 4	T 3	T 14	V 6		S 3	
			G 3	Y 1	I 2	A 5	P 1	Q 4	A 3	V 2	S 10	IQR 9		Q 2	
			H 1	D 1	T 2	R 3	Q 1	A 1	K 3	G 1	P 10	L 2			
				I 1	DK 1	E 1	E 2	IR 1		F 1	F 5	K 2			
				Q 1	H 1	P 1	G 1	RQ 1		Y 1	N 4	F 1			
					K 1		T 1			E 1	A 2	S 1			
											K 1	G 1			
											G 1	Y 1			
											V 1	SRG 1			
												YQR 1			

y_1	y_2	y_3	y_4	y_5	y_6	y_7	y_8	y_9	y_{10}	y_{11}	y_{12}	y_{13}	y_{14}	y_{15}	y_{16}	y_{17}
R 121	A 104	F 75	Y 94	T 81	T 102	G 78	R 34	I 115	I 105	G 130	D 106	I 119	R 131	Q 115	A 137	H 116
K 9	V 14	I 25	V 13	A 44	R 11	E 17	K 23	V 7	T 12	R 3	N 23	M 16	G 2	K 16		Y 8
Q 5	N 5	V 15	H 14	V 4	I 11	K 14	Q 20	R 2	V 7	E 1	I 3	ME 1	K 1	R 6		R 8
G 1	R 4	L 8	L 8	Q 3	Q 5	R 10	E 20	N 1	K 2	K 1	R 3	L 1	S 1			Q 3
S 1	K 4	W 5	F 1	H 1	A 5	T 3	D 13	G 1	M 2	A 1	T 1		E 1			
T 1	T 2	Y 5	S 2	I 1	M 1	D 5	N 9		R 2		S 1					
	S 2	G 1	I 1	S 1		Q 4	A 8		L 1							
	F 1	S 1	T 1	Y 2		A 3	G 4		S 1							
	P 1	T 1	M 1			H 1	S 2		E 1							
	W 1		R 1			N 1	I 1		Q 1							
			VH 1			P 1			A 1							
			FT 1													

Table 11.

[illegible] # NAME	TRRRIEIGPGRAFYTTG	TYPE(#/TOTAL)	MISMATCHES (1 TO 13)
[illegible]	[illegible]	[illegible]	[illegible]

Table 12

COMMON SEQUENCE PATTERNS

SEQUENCE	APPROXIMATE OCCURENCE (%)
I a I G P G R	60
I a I G P G R A	50
I a I G P G R A F	32
I G P G R A	52
G P G R A F	40
I G P G R A F	35
I G P G R A a Y	31

Table 13

IIIB
MetThrArgIleGlnArgGlyProGlyArgAlaPheVal GlyGlyGlyGly -

RF
SerIleThrArgGlyProGlyArgValIleTyr GlyGlyGlyGly -

MN
ArgIleHisIleGlyProGlyArgAlaPheTyr GlyGlyGlyGly -

SC
SerIleHisIleGlyProGlyArgAlaPheTyr GlyGlyGlyGly -

WMJ1
HisIleHisIleGlyProGlyArgAlaPheTyr GlyGlyGlyGly -

LeuSerIleCys

Table 13A

E.coli BG

MetLeuArgProValGluThrProThrArgGluIleLysLysLeuAspGlyLeuTrp -

AlaPheSerLeuAspArgGluArgValValArgTyrHisArgTrpIleArgGlnAlaSer -

IIIB
MetThrArgIleGlnArgGlyProGlyArgAlaPheVal GlyGlyGlyGly -

RF
SerIleThrArgGlyProGlyArgValIleTyr GlyGlyGlyGly -

MN
ArgIleHisIleGlyProGlyArgAlaPheTyr GlyGlyGlyGly -

SC
SerIleHisIleGlyProGlyArgAlaPheTyr GlyGlyGlyGly -

WMJ1
HisIleHisIleGlyProGlyArgAlaPheTyr GlyGlyGlyGly -

LeuSerIleCys

Claims

1. A compound having the capability of eliciting, and/or binding with, neutralizing antibodies where said capability results from an amino acid sequence which:

(a) is the principal neutralizing domain of an HIV variant;

(b) is a portion of the principal neutralizing domain of an HIV variant; or

(c) is equivalent either to a principal neutralizing domain or a portion thereof.

2. A compound, other than naturally occurring HIV envelope protein, said compound having the capability of eliciting, and/or binding with, neutralizing antibodies, said compound comprising the principal neutralizing domain, or a segment thereof, of an HIV variant.

3. The compound, according to claim 2, wherein said compound is modified by addition of one or more of the following moieties: cysteine, a protein or other moiety capable of enhancing immunogenicity, a peptide from an HIV principal neutralizing domain, a peptide capable of stimulating T-cells, or a general immune stimulant.

4. A compound having the capability of eliciting, and/or binding with, neutralizing antibodies, said compound consisting essentially of the amino acids between cysteine residues occurring at, or around, positions 296 and 331 of the envelope protein of an HIV variant.

5. A compound having the capability of eliciting, and/or binding with, HIV neutralizing antibodies, said compound having the formula

$$a\ x\ G\ z\ G\ y\ b$$

wherein x is 0 to 13 amino acids in length;
y is 0 to 17 amino acids in length; and
z is P, A, S, Q, or L; and

either a or b, but not both, may be omitted; either a or b individually may comprise any one of the following: cysteine, a protein or other moiety capable of enhancing immunogenicity, a peptide from an HIV principal neutralizing domain, a peptide capable of stimulating T-cells, or a general immune stimulant.

6. The compound, according to claim 5, wherein said compound is circularized.

7. The compound, according to claim 5, wherein said compound comprises epitopes from the principal neutralizing domain of more than one HIV variant.

8. A polypeptide selected from the group consisting of
(1) HIV 10 Kd fusion protein denoted Sub 1;
(2) HIV protein portion of Sub 1;
(3) HIV 18 Kd fusion protein denoted Sub 2;
(4) HIV protein portion of Sub 2;
(5) HIV 27 Kd fusion protein denoted $PB1_{RF}$;
(6) HIV protein portion of $PB1_{RF}$;
(7) HIV 28 Kd fusion protein denoted $PB1_{MN}$;
(8) HIV protein portion of $PB1_{MN}$;
(9) HIV 26 Kd fusion protein denoted $PB1_{SC}$;
(10) HIV protein portion of $PB1_{SC}$;
(11) HIV 26 Kd fusion protein denoted $PB1_{WMJ2}$;
(12) HIV protein portion of $PB1_{WMJ2}$;
(13) peptide III_B(BH10)-PND;
(14) peptide RF-PND;
(15) peptide MN-PND;
(16) peptide SC-PND;
(17) peptide WMJ-2-PND;
(18) peptide LAV-MAL-PND;
(19) peptide SF-2-PND;
(20) peptide NY5-PND;
(21) peptide Z3-PND;
(22) peptide WMJ1-PND;
(23) peptide WMJ3-PND;
(24) peptide Z6-PND;
(25) peptide LAVELI-PND;
(26) peptide CDC451-PND;
(27) peptide CDC42-PND;
(28) peptide BAL-PND;
(29) peptide HIV-2-PND;
(30) peptide 135;
(31) peptide 136;
(32) peptide 139;
(33) peptide 141;
(34) peptide 142;
(35) peptide 143;
(36) peptide 131;

38	(37)	peptide 132;
39	(38)	peptide 134;
40	(39)	peptide 339;
41	(40)	RP342 (WMJ2);
42	(41)	RP343 (SC);
43	(42)	RP60 ($\mathrm{III_B}$);
44	(43)	RP335 ($\mathrm{III_B}$);
45	(44)	RP337 ($\mathrm{III_B}$);
46	(45)	RP77 ($\mathrm{III_B}$);
47	(46)	RP83 (WMJ1);
48	(47)	RP79 ($\mathrm{III_B}$);
49	(48)	RP57;
50	(49)	RP55;
51	(50)	RP75;
52	(51)	RP56;
53	(52)	RP59;
54	(53)	RP73 ($\mathrm{III_B}$,RF);
55	(54)	RP74 ($\mathrm{III_B}$,RF,MN,SC);
56	(55)	RP80 ($\mathrm{III_B}$,RF);
57	(56)	RP81 ($\mathrm{III_B}$,RF,WMJ1,MN);
58	(57)	RP82 (WMJ1,MN);
59	(58)	RP137 ($\mathrm{III_B}$,RF);
60	(59)	RP140 ($\mathrm{III_B}$,RF);
61	(60)	peptide 64 (HIV-IIIB/HIV-RF/HIV-MN/HIV-SC);
62	(61)	peptide 338 (HIV-$\mathrm{III_B}$/HIV-RF);
63	(62)	peptide 138;
64	(63)	RP342;
65	(64)	RP96;
66	(65)	RP97;
67	(66)	RP98;
68	(67)	RP99;
69	(68)	RP100;
70	(69)	RP102;
71	(70)	RP88;
72	(71)	RP91;
73	(72)	RP104;
74	(73)	RP106;

(74) RP108;
(75) RP70;
(76) RP84;
(77) RP144;
(78) RP145;
(79) RP146;
(80) RP147;
(81) RP150;
(82) RP151;
(83) RP63;
(84) RP41;
(85) RP61;
(86) RP75;
(87) RP111;
(88) RP113;
(89) RP114;
(90) RP116;
(91) RP120;
(92) RP121c;
(93) RP122c;
(94) RP123c;

and chemical modifications thereof.

9. A DNA sequence which codes for a polypeptide, other than the naturally occurring HIV envelope protein, said polypeptide having the capability of eliciting, and/or binding with, neutralizing antibodies, said polypeptide comprising the principal neutralizing domain, or a segment thereof, of an HIV variant.

10. The DNA sequence, according to claim 9, wherein said polypeptide has the formula

$$a\ x\ G\ z\ G\ y\ b$$

wherein x is 0 to 13 amino acids in length;
y is 0 to 17 amino acids in length; and
z is P, A, S, Q, or L; and

either a or b, but not both, may be omitted; either a or b individually may comprise any one of the following: cysteine, a protein or other moiety capable of enhancing immunogenicity, a peptide from an HIV principal neutralizing domain, a peptide capable of stimulating T-cells, or a general immune stimulant.

11. A composition capable of eliciting and/or binding with neutralizing antibodies to a broad range of HIV variants, said composition comprising one or more compounds having the following formula:

$$a\ x\ G\ z\ G\ y\ b$$

wherein x is 0 to 13 amino acids in length;
y is 0 to 17 amino acids in length; and
z is P, A, S, Q, or L; and

either a or b, but not both, may be omitted; either a or b individually may comprise any one of the following: cysteine, a protein or other moiety capable of enhancing immunogenicity, a peptide from an HIV principal neutralizing domain, a peptide capable of stimulating T-cells, or a general immune stimulant.

12. A composition, according to claim 11, wherein x is selected from:

$x(0) \equiv x$ is not present
$x(1) \equiv x_1$
$x(2) \equiv x_2\ x_1$
$x(3) \equiv x_3\ x_2\ x_1$
$x(4) \equiv x_4\ x_3\ x_2\ x_1$
$x(5) \equiv x_5\ x_4\ x_3\ x_2\ x_1$
$x(6) \equiv x_6\ x_5\ x_4\ x_3\ x_2\ x_1$
$x(7) \equiv x_7\ x_6\ x_5\ x_4\ x_3\ x_2\ x_1$
$x(8) \equiv x_8\ x_7\ x_6\ x_5\ x_4\ x_3\ x_2\ x_1$
$x(9) \equiv x_9\ x_8\ x_7\ x_6\ x_5\ x_4\ x_3\ x_2\ x_1$
$x(10) \equiv x_{10}\ x_9\ x_8\ x_7\ x_6\ x_5\ x_4\ x_3\ x_2\ x_1$
$x(11) \equiv x_{11}\ x_{10}\ x_9\ x_8\ x_7\ x_6\ x_5\ x_4\ x_3\ x_2\ x_1$
$x(12) \equiv x_{12}\ x_{11}\ x_{10}\ x_9\ x_8\ x_7\ x_6\ x_5\ x_4\ x_3\ x_2\ x_1$
$x(13) \equiv x_{13}\ x_{12}\ x_{11}\ x_{10}\ x_9\ x_8\ x_7\ x_6\ x_5\ x_4\ x_3\ x_2\ x_1$
z is P, L, A, S, or Q; and y is selected from:
$y(0) \equiv y$ is not present
$y(1) \equiv y_1$
$y(2) \equiv y_1\ y_2$
$y(3) \equiv y_1\ y_2\ y_3$
$y(4) \equiv y_1\ y_2\ y_3\ y_4$
$y(5) \equiv y_1\ y_2\ y_3\ y_4\ y_5$
$y(6) \equiv y_1\ y_2\ y_3\ y_4\ y_5\ y_6$
$y(7) \equiv y_1\ y_2\ y_3\ y_4\ y_5\ y_6\ y_7$
$y(8) \equiv y_1\ y_2\ y_3\ y_4\ y_5\ y_6\ y_7\ y_8$

$y(9) = y_1 y_2 y_3 y_4 y_5 y_6 y_7 y_8 y_9$

$y(10) = y_1 y_2 y_3 y_4 y_5 y_6 y_7 y_8 y_9 y_{10}$

$y(11) = y_1 y_2 y_3 y_4 y_5 y_6 y_7 y_8 y_9 y_{10} y_{11}$

$y(12) = y_1 y_2 y_3 y_4 y_5 y_6 y_7 y_8 y_9 y_{10} y_{11} y_{12}$

$y(13) = y_1 y_2 y_3 y_4 y_5 y_6 y_7 y_8 y_9 y_{10} y_{11} y_{12} y_{13}$

$y(14) = y_1 y_2 y_3 y_4 y_5 y_6 y_7 y_8 y_9 y_{10} y_{11} y_{12} y_{13} y_{14}$

$y(15) = y_1 y_2 y_3 y_4 y_5 y_6 y_7 y_8 y_9 y_{10} y_{11} y_{12} y_{13} y_{14} y_{15}$

$y(16) = y_1 y_2 y_3 y_4 y_5 y_6 y_7 y_8 y_9 y_{10} y_{11} y_{12} y_{13} y_{14} y_{15} y_{16}$

$y(17) = y_1 y_2 y_3 y_4 y_5 y_6 y_7 y_8 y_9 y_{10} y_{11} y_{12} y_{13} y_{14} y_{15} y_{16} y_{17}$

and wherein

x_1 is I, R, M, IQR, V, L, K, F, S, G, Y, SRG, or YQR;

x_2 is H, R, Y, T, S, P, F, N, A, K, G, or V;

x_3 is I, L, M, T, V, E, G, F, or Y;

x_4 is R, S, G, H, A, K, or not present;

x_5 is K, R, I, N, Q, A, IR, RQ, or not present;

x_6 is R, K, S, I, P, Q, E, G, or T;

x_7 is T, K, V, I, A, R, P, or E;

x_8 is N, NV, Y, KI, I, T, DK, H, or K;

x_9 is N, S, K, E, Y, D, I, or Q;

x_{10} is N, Y, S, D, G, or H;

x_{11} is P;

x_{12} is R, I, or K;

x_{13} is T, I, M or A;

y_1 is R, K, Q, G, S, or T;

y_2 is A, V, N, R, K, T, S, F, P, or W;

y_3 is F, I, V, L, W, Y, G, S, or T;

y_4 is Y, V, H, L, F, S, I, T, M, R, VH, or FT;

y_5 is T, A, V, Q, H, I, S, Y, or not present;

y_6 is T, R, I, Q, A, M, or not present;

y_7 is G, E, K, R, T, D, Q, A, H, N, P, or not present;

y_8 is R, Q, E, K, D, N, A, G, S, I, or not present;

y_9 is I, V, R, N, G, or not present;

y_{10} is I, T, V, K, M, R, L, S, E, Q, A, or not present;

y_{11} is G, R, E, K, H, or not present;

y_{12} is D, N, I, R, T, S, or not present;

y_{13} is I, M, ME, L, or not present;

y_{14} is R, G, K, S, E, or not present;

y_{15} is Q, K, or R;
y_{16} is A; and
y_{17} is H, Y, R, or Q.

13. The composition, according to claim 12, wherein
x_1 is I;
z is P;
y_1 is R; and
y_2 is A.

14. The composition, according to claim 12, wherein
x_1 is I;
x_3 is I;
z is P; and
y_1 is R.

15. The composition, according to claim 12, wherein
z is P;
y_1 is R;
y_2 is A; and
y_3 is F.

16. A composition, according to claim 12, wherein
x_1 is I;
x_2 is H;
x_3 is I;
x_4 is R;
x_5 is K;
x_6 is R;
x_7 is T;
z is P;
y_1 is R;
y_2 is A;
y_3 is F;
y_4 is Y;
y_5 is T;
y_6 is T; and

y_7 is G.

17. The composition, according to claim 11, wherein said compound is circularized.

18. The composition, according to claim 11, wherein a and/or b comprise a peptide from an HIV principal neutralizing domain.

19. The composition, according to claim 11, wherein said moiety capable of enhancing immunogenicity is a viral particle, microorganism, or immunogenic portion thereof.

20. A prophylactic or therapeutic composition comprising immune globulin, monoclonal antibodies, and/or polyclonal antibodies generated by immunizing an appropriate animal such as a mouse, rat, horse, goat, human, or chimpanzee with a hybrid compound which comprises compounds where said polypeptides are not naturally occurring HIV envelope proteins, said polypeptides having the capability of eliciting, and/or binding with, neutralizing antibodies, said polypeptides comprising the principal neutralizing domain, or a segment thereof, of an HIV variant.

21. A prophylactic or therapeutic composition comprising antibodies generated by immunizing an animal or human with at least one compound comprising a principal neutralizing domain of an HIV variant followed by immunization of said animal or human with

(a) a mixture comprising compounds, where said compounds are not naturally occurring HIV envelope proteins, said compounds having the capability of eliciting, and/or binding with, neutralizing antibodies, said compounds comprising the principal neutralizing domain, or a segment thereof, of an HIV variant; or

(b) a hybrid compound which comprises polypeptides, where said polypeptides are not naturally occurring HIV envelope proteins, said polypeptides having the capability of eliciting, and/or binding with, neutralizing antibodies, said polypeptides comprising the principal neutralizing domain, or a segment thereof, of an HIV variant.

22. A prophylactic or therapeutic composition, comprising an antibody raised against a mixture comprising compounds, where said compounds are not naturally occurring HIV envelope proteins, but have the capability of eliciting, and/or binding with, neutralizing antibodies, each of said compounds comprising the principal neutralizing domain, or a segment thereof, of an HIV variant.

23. A prophylactic or therapeutic composition, comprising an antibody raised against at least one compound having the capability of eliciting, and/or binding with, HIV neutralizing antibodies, said compound having the formula

a x G z G y b

wherein **x** is 0 to 13 amino acids in length;
y is 0 to 17 amino acids in length; and
z is P, A, S, Q, or L; and

either **a** or **b**, but not both, may be omitted; either **a** or **b** individually may comprise any one of the following: cysteine, a protein or other moiety capable of enhancing immunogenicity, a peptide from an HIV principal neutralizing domain, a peptide capable of stimulating T-cells, or a general immune stimulant.

24. The composition, according to claim 23, wherein **a** and/or **b** comprises a peptide from an HIV principal neutralizing domain.

25. An antibody raised against a compound, said compound having the formula

a x G z G y b

wherein **x** is 0 to 13 amino acids in length;
y is 0 to 17 amino acids in length; and
z is P, A, S, Q, or L; and

either **a** or **b**, but not both, may be omitted; either **a** or **b** individually may comprise any one of the following: cysteine, a protein or other moiety capable of enhancing immunogenicity, a peptide from an HIV principal neutralizing domain, a peptide capable of stimulating T-cells, or a general immune stimulant.

26. The antibody, according to claim 25, wherein **x** is selected from

$x(0) \equiv$ x is not present
$x(1) \equiv x_1$
$x(2) \equiv x_2\ x_1$
$x(3) \equiv x_3\ x_2\ x_1$
$x(4) \equiv x_4\ x_3\ x_2\ x_1$
$x(5) \equiv x_5\ x_4\ x_3\ x_2\ x_1$
$x(6) \equiv x_6\ x_5\ x_4\ x_3\ x_2\ x_1$
$x(7) \equiv x_7\ x_6\ x_5\ x_4\ x_3\ x_2\ x_1$
$x(8) \equiv x_8\ x_7\ x_6\ x_5\ x_4\ x_3\ x_2\ x_1$
$x(9) \equiv x_9\ x_8\ x_7\ x_6\ x_5\ x_4\ x_3\ x_2\ x_1$
$x(10) \equiv x_{10}\ x_9\ x_8\ x_7\ x_6\ x_5\ x_4\ x_3\ x_2\ x_1$

$x(11) \equiv x_{11}\ x_{10}\ x_9\ x_8\ x_7\ x_6\ x_5\ x_4\ x_3\ x_2\ x_1$

$x(12) \equiv x_{12}\ x_{11}\ x_{10}\ x_9\ x_8\ x_7\ x_6\ x_5\ x_4\ x_3\ x_2\ x_1$

$x(13) \equiv x_{13}\ x_{12}\ x_{11}\ x_{10}\ x_9\ x_8\ x_7\ x_6\ x_5\ x_4\ x_3\ x_2\ x_1$

z is P, L, A, S, or Q; and y is selected from:

$y(0) \equiv$ y is not present

$y(1) \equiv y_1$

$y(2) \equiv y_1\ y_2$

$y(3) \equiv y_1\ y_2\ y_3$

$y(4) \equiv y_1\ y_2\ y_3\ y_4$

$y(5) \equiv y_1\ y_2\ y_3\ y_4\ y_5$

$y(6) \equiv y_1\ y_2\ y_3\ y_4\ y_5\ y_6$

$y(7) \equiv y_1\ y_2\ y_3\ y_4\ y_5\ y_6\ y_7$

$y(8) \equiv y_1\ y_2\ y_3\ y_4\ y_5\ y_6\ y_7\ y_8$

$y(9) \equiv y_1\ y_2\ y_3\ y_4\ y_5\ y_6\ y_7\ y_8\ y_9$

$y(10) \equiv y_1\ y_2\ y_3\ y_4\ y_5\ y_6\ y_7\ y_8\ y_9\ y_{10}$

$y(11) \equiv y_1\ y_2\ y_3\ y_4\ y_5\ y_6\ y_7\ y_8\ y_9\ y_{10}\ y_{11}$

$y(12) \equiv y_1\ y_2\ y_3\ y_4\ y_5\ y_6\ y_7\ y_8\ y_9\ y_{10}\ y_{11}\ y_{12}$

$y(13) \equiv y_1\ y_2\ y_3\ y_4\ y_5\ y_6\ y_7\ y_8\ y_9\ y_{10}\ y_{11}\ y_{12}\ y_{13}$

$y(14) \equiv y_1\ y_2\ y_3\ y_4\ y_5\ y_6\ y_7\ y_8\ y_9\ y_{10}\ y_{11}\ y_{12}\ y_{13}\ y_{14}$

$y(15) \equiv y_1\ y_2\ y_3\ y_4\ y_5\ y_6\ y_7\ y_8\ y_9\ y_{10}\ y_{11}\ y_{12}\ y_{13}\ y_{14}\ y_{15}$

$y(16) \equiv y_1\ y_2\ y_3\ y_4\ y_5\ y_6\ y_7\ y_8\ y_9\ y_{10}\ y_{11}\ y_{12}\ y_{13}\ y_{14}\ y_{15}\ y_{16}$

$y(17) \equiv y_1\ y_2\ y_3\ y_4\ y_5\ y_6\ y_7\ y_8\ y_9\ y_{10}\ y_{11}\ y_{12}\ y_{13}\ y_{14}\ y_{15}\ y_{16}\ y_{17}$

and wherein

x_1 is I, R, M, IQR, V, L, K, F, S, G, Y, SRG, or YQR;

x_2 is H, R, Y, T, S, P, F, N, A, K, G, or V;

x_3 is I, L, M, T, V, E, G, F, or Y;

x_4 is R, S, G, H, A, K, or not present;

x_5 is K, R, I, N, Q, A, IR, RQ, or not present;

x_6 is R, K, S, I, P, Q, E, G, or T;

x_7 is T, K, V, I, A, R, P, or E;

x_8 is N, NV, Y, KI, I, T, DK, H, or K;

x_9 is N, S, K, E, Y, D, I, or Q;

x_{10} is N, Y, S, D, G, or H;

x_{11} is P;

x_{12} is R, I, or K;

x_{13} is T, I, M or A;

y_1 is R, K, Q, G, S, or T;

y_2 is A, V, N, R, K, T, S, F, P, or W;
y_3 is F, I, V, L, W, Y, G, S, or T;
y_4 is Y, V, H, L, F, S, I, T, M, R, VH, or FT;
y_5 is T, A, V, Q, H, I, S, Y, or not present;
y_6 is T, R, I, Q, A, M, or not present;
y_7 is G, E, K, R, T, D, Q, A, H, N, P, or not present;
y_8 is R, Q, E, K, D, N, A, G, S, I, or not present;
y_9 is I, V, R, N, G, or not present;
y_{10} is I, T, V, K, M, R, L, S, E, Q, A, or not present;
y_{11} is G, R, E, K, H, or not present;
y_{12} is D, N, I, R, T, S, or not present;
y_{13} is I, M, ME, L, or not present;
y_{14} is R, G, K, S, E, or not present;
y_{15} is Q, K, or R;
y_{16} is A; and
y_{17} is H, Y, R, or Q.

27. The antibody, according to claim 26, wherein said compound(s) are selected from the group consisting of
(a) a-Y-S-N-V-R-N-R-I-H-I-G-P-G-R-A-F-H-T-T-K-R-I-T-b;
(b) a-N-N-N-T-S-R-G-I-R-I-G-P-G-R-A-I-L-A-T-E-R-I-I-b;
(c) a-N-N-N-T-R-K-G-I-F-I-G-P-G-R-N-I-Y-T-T-G-N-I-I-b;
(d) a-N-T-R-K-S-I-R-I-Q-R-G-P-G-R-A-F-V-T-I-G-K-I-G-b;
(e) a-N-N-N-T-R-K-R-(I or V)-T-M-G-P-G-R-V-(Y or W)-Y-(X or T)-(A or T)-G-Q-I-I-b;
(f) a-N-N-N-(I or T)-R-K-(R or S)-I-T-(R or K)-G-P-G-(R or K)-V-I-Y-A-T-G-Q-I-I-b;
(g) a-N-N-N-T-R-K-G-I-Y-V-G-S-G-R-(A or K)-V-T-T-R-(D or H or Q)-K-I-(I or M)-b; and
(h) a-T-R-Q-(R or S)-T-P-I-G-L-G-Q-(A or S)-L-Y-T-T-R-b.

28. A vaccine composition for generating a broadly neutralizing immunological response, said vaccine composition comprising a mixture comprising compounds, where said compounds are not naturally occurring HIV envelope proteins, but have the capability of eliciting, and/or binding with, neutralizing antibodies, said compounds comprising the principal neutralizing domain, or a segment thereof, of HIV variants.

29. A vaccine composition, for generating a broadly neutralizing immunological response, said vaccine composition comprising a hybrid compound which comprises polypeptides, where said polypeptides are not naturally occurring HIV envelope proteins, but have the capability of eliciting,

and/or binding with, neutralizing antibodies, said polypeptides comprising the principal neutralizing domain, or a segment thereof, of an HIV variant.

30. A vaccine composition comprising at least one compound having the capability of eliciting, and/or binding with, HIV neutralizing antibodies, said compound having the formula

a x G z G y b

wherein x is 0 to 13 amino acids in length;

y is 0 to 17 amino acids in length; and

z is P, A, S, Q, or L; and

either a or b, but not both, may be omitted; either a or b individually may comprise any one of the following: cysteine, a protein or other moiety capable of enhancing immunogenicity, a peptide from an HIV principal neutralizing domain, a peptide capable of stimulating T-cells, or a general immune stimulant.

31. The vaccine composition, according to claim 30, wherein said composition is formulated with an immunological adjuvant.

32. The vaccine, according to claim 30, wherein x is selected from:

$x(0) \equiv$ x is not present

$x(1) \equiv x_1$

$x(2) \equiv x_2\ x_1$

$x(3) \equiv x_3\ x_2\ x_1$

$x(4) \equiv x_4\ x_3\ x_2\ x_1$

$x(5) \equiv x_5\ x_4\ x_3\ x_2\ x_1$

$x(6) \equiv x_6\ x_5\ x_4\ x_3\ x_2\ x_1$

$x(7) \equiv x_7\ x_6\ x_5\ x_4\ x_3\ x_2\ x_1$

$x(8) \equiv x_8\ x_7\ x_6\ x_5\ x_4\ x_3\ x_2\ x_1$

$x(9) \equiv x_9\ x_8\ x_7\ x_6\ x_5\ x_4\ x_3\ x_2\ x_1$

$x(10) \equiv x_{10}\ x_9\ x_8\ x_7\ x_6\ x_5\ x_4\ x_3\ x_2\ x_1$

$x(11) \equiv x_{11}\ x_{10}\ x_9\ x_8\ x_7\ x_6\ x_5\ x_4\ x_3\ x_2\ x_1$

$x(12) \equiv x_{12}\ x_{11}\ x_{10}\ x_9\ x_8\ x_7\ x_6\ x_5\ x_4\ x_3\ x_2\ x_1$

$x(13) \equiv x_{13}\ x_{12}\ x_{11}\ x_{10}\ x_9\ x_8\ x_7\ x_6\ x_5\ x_4\ x_3\ x_2\ x_1$

z is P, L, A, S, or Q; and y is selected from:

$y(0) \equiv$ y is not present

$y(1) \equiv y_1$

$y(2) \equiv y_1\ y_2$

$y(3) \equiv y_1\ y_2\ y_3$

$y(4) \equiv y_1\, y_2\, y_3\, y_4$

$y(5) \equiv y_1\, y_2\, y_3\, y_4\, y_5$

$y(6) \equiv y_1\, y_2\, y_3\, y_4\, y_5\, y_6$

$y(7) \equiv y_1\, y_2\, y_3\, y_4\, y_5\, y_6\, y_7$

$y(8) \equiv y_1\, y_2\, y_3\, y_4\, y_5\, y_6\, y_7\, y_8$

$y(9) \equiv y_1\, y_2\, y_3\, y_4\, y_5\, y_6\, y_7\, y_8\, y_9$

$y(10) \equiv y_1\, y_2\, y_3\, y_4\, y_5\, y_6\, y_7\, y_8\, y_9\, y_{10}$

$y(11) \equiv y_1\, y_2\, y_3\, y_4\, y_5\, y_6\, y_7\, y_8\, y_9\, y_{10}\, y_{11}$

$y(12) \equiv y_1\, y_2\, y_3\, y_4\, y_5\, y_6\, y_7\, y_8\, y_9\, y_{10}\, y_{11}\, y_{12}$

$y(13) \equiv y_1\, y_2\, y_3\, y_4\, y_5\, y_6\, y_7\, y_8\, y_9\, y_{10}\, y_{11}\, y_{12}\, y_{13}$

$y(14) \equiv y_1\, y_2\, y_3\, y_4\, y_5\, y_6\, y_7\, y_8\, y_9\, y_{10}\, y_{11}\, y_{12}\, y_{13}\, y_{14}$

$y(15) \equiv y_1\, y_2\, y_3\, y_4\, y_5\, y_6\, y_7\, y_8\, y_9\, y_{10}\, y_{11}\, y_{12}\, y_{13}\, y_{14}\, y_{15}$

$y(16) \equiv y_1\, y_2\, y_3\, y_4\, y_5\, y_6\, y_7\, y_8\, y_9\, y_{10}\, y_{11}\, y_{12}\, y_{13}\, y_{14}\, y_{15}\, y_{16}$

$y(17) \equiv y_1\, y_2\, y_3\, y_4\, y_5\, y_6\, y_7\, y_8\, y_9\, y_{10}\, y_{11}\, y_{12}\, y_{13}\, y_{14}\, y_{15}\, y_{16}\, y_{17}$

and wherein

x_1 is I, R, M, IQR, V, L, K, F, S, G, Y, SRG, or YQR;

x_2 is H, R, Y, T, S, P, F, N, A, K, G, or V;

x_3 is I, L, M, T, V, E, G, F, or Y;

x_4 is R, S, G, H, A, K, or not present;

x_5 is K, R, I, N, Q, A, IR, RQ, or not present;

x_6 is R, K, S, I, P, Q, E, G, or T;

x_7 is T, K, V, I, A, R, P, or E;

x_8 is N, NV, Y, KI, I, T, DK, H, or K;

x_9 is N, S, K, E, Y, D, I, or Q;

x_{10} is N, Y, S, D, G, or H;

x_{11} is P;

x_{12} is R, I, or K;

x_{13} is T, I, M or A;

y_1 is R, K, Q, G, S, or T;

y_2 is A, V, N, R, K, T, S, F, P, or W;

y_3 is F, I, V, L, W, Y, G, S, or T;

y_4 is Y, V, H, L, F, S, I, T, M, R, VH, or FT;

y_5 is T, A, V, Q, H, I, S, Y, or not present;

y_6 is T, R, I, Q, A, M, or not present;

y_7 is G, E, K, R, T, D, Q, A, H, N, P, or not present;

y_8 is R, Q, E, K, D, N, A, G, S, I, or not present;

y_9 is I, V, R, N, G, or not present;

y_{10} is I, T, V, K, M, R, L, S, E, Q, A, or not present;
y_{11} is G, R, E, K, H, or not present;
y_{12} is D, N, I, R, T, S, or not present;
y_{13} is I, M, ME, L, or not present;
y_{14} is R, G, K, S, E, or not present;
y_{15} is Q, K, or R;
y_{16} is A; and
y_{17} is H, Y, R, or Q.

33. The vaccine, according to claim 32, wherein
x_1 is I;
x_2 is H;
x_3 is I;
x_4 is R;
x_5 is K;
x_6 is R;
x_7 is T;
z is P;
y_1 is R;
y_2 is A;
y_3 is F;
y_4 is Y;
y_5 is T;
y_6 is T; and
y_7 is G.

34. The vaccine, according to claim 32, wherein said compound(s) are selected from the group consisting of
(a) a-Y-S-N-V-R-N-R-I-H-I-G-P-G-R-A-F-H-T-T-K-R-I-T-b;
(b) a-N-N-N-T-S-R-G-I-R-I-G-P-G-R-A-I-L-A-T-E-R-I-I-b;
(c) a-N-N-N-T-R-K-G-I-F-I-G-P-G-R-N-I-Y-T-T-G-N-I-I-b;
(d) a-N-T-R-K-S-I-R-I-Q-R-G-P-G-R-A-F-V-T-I-G-K-I-G-b;
(e) a-N-N-N-T-R-K-R-(I or V)-T-M-G-P-G-R-V-(Y or W)-Y-(X or T)-(A or T)-G-Q-I-I-b;
(f) a-N-N-N-(I or T)-R-K-(R or S)-I-T-(R or K)-G-P-G-(R or K)-V-I-Y-A-T-G-Q-I-I-b;
(g) a-N-N-N-T-R-K-G-I-Y-V-G-S-G-R-(A or K)-V-T-T-R-(D or H or Q)-K-I-(I or M)-b; and
(h) a-T-R-Q-(R or S)-T-P-I-G-L-G-Q-(A or S)-L-Y-T-T-R-b.

35. The vaccine composition of claim 32 wherein said compound(s) are capable of eliciting antibodies that bind to the sequence G-P-G-R-A-F.

36. The vaccine composition of claim 32 wherein said compound(s) are capable of eliciting antibodies that bind to the sequence I-G-P-G-R-A-F.

37. The vaccine composition of claim 32 wherein said compound(s) are capable of eliciting antibodies that bind to the sequence I-G-P-G-R-A.

38. The vaccine composition of claim 32 wherein said compound(s) are capable of eliciting antibodies that bind to the sequence I-a-I-G-P-G-R, wherein a is any of the 20 amino acids.

39. The vaccine, according to claim 38, wherein a is H.

40. The vaccine composition of claim 32 wherein said compound(s) are capable of eliciting antibodies that bind to the sequence I-a-I-G-P-G-R-A, wherein a is any of the 20 amino acids.

41. The vaccine, according to claim 40, wherein a is H.

42. The vaccine composition of claim 32 wherein said compound(s) are capable of eliciting antibodies that bind to the sequence I-a-I-G-P-G-R-A-F, wherein a is any of the 20 amino acids.

43. The vaccine, according to claim 42, wherein a is H.

44. The vaccine composition, according to claim 30, wherein said compound(s) are circularized.

45. The vaccine composition, according to claim 30, wherein said compound(s) comprise epitopes from more than one HIV variant.

46. The vaccine, according to claim 32, wherein

x_1 is I;

z is P;

y_1 is R; and

y_2 is A.

47. The composition, according to claim 32, wherein

x_1 is I;

x_3 is I;

z is P; and

y_1 is R.

48. The composition, according to claim 32, wherein

z is P;

y_1 is R;

y_2 is A; and

y_3 is F.

49. The vaccine, according to claim 32, wherein x is selected from the group consisting of x(5), x(6), x(7), x(8), x(9), x(10), and x(11); and y is selected from y(5), y(6), y(7), y(8), y(9), y(10), and y(11).

50. The composition, according to claim 32, wherein said compound has the following amino acid sequence:

$a\text{-}x_{13}\text{-}R\text{-}P\text{-}x_{10}\text{-}x_9\text{-}x_8\text{-}x_7\text{-}x_6\text{-}x_5\text{-}x_4\text{-}x_3\text{-}x_2\text{-}x_1\text{-}G\text{-}z\text{-}G\text{-}y_1\text{-}y_2\text{-}y_3\text{-}y_4\text{-}y_5\text{-}y_6\text{-}y_7\text{-}y_8\text{-}y_9\text{-}y_{10}\text{-}G\text{-}y_{12}\text{-}y_{13}\text{-}R\text{-}y_{15}\text{-}A\text{-}y_{17}\text{-}b$.

51. A kit for use in detecting antibody against HIV in a biological fluid, said kit comprising:

(a) a compound, where said compound is not a naturally occurring HIV envelope protein, said compound having the capability of eliciting, and/or binding with, neutralizing antibodies, said compound comprising the principal neutralizing domain, or a segment thereof, of an HIV variant; and

(b) a means for detecting complexes formed between said antibody and said compound.

52. The kit, according to claim 51, wherein said compound has the formula

x G z G y

wherein x is 0 to 13 amino acids in length;

y is 0 to 17 amino acids in length; and

z is P, A, S, Q, or L.

53. The kit, according to claim 52, wherein x is selected from

x(0) ≡ x is not present

x(1) ≡ x_1

x(2) ≡ $x_2\ x_1$

x(3) ≡ $x_3\ x_2\ x_1$

x(4) ≡ $x_4\ x_3\ x_2\ x_1$

x(5) ≡ $x_5\ x_4\ x_3\ x_2\ x_1$

$x(6) \equiv x_6\, x_5\, x_4\, x_3\, x_2\, x_1$

$x(7) \equiv x_7\, x_6\, x_5\, x_4\, x_3\, x_2\, x_1$

$x(8) \equiv x_8\, x_7\, x_6\, x_5\, x_4\, x_3\, x_2\, x_1$

$x(9) \equiv x_9\, x_8\, x_7\, x_6\, x_5\, x_4\, x_3\, x_2\, x_1$

$x(10) \equiv x_{10}\, x_9\, x_8\, x_7\, x_6\, x_5\, x_4\, x_3\, x_2\, x_1$

$x(11) \equiv x_{11}\, x_{10}\, x_9\, x_8\, x_7\, x_6\, x_5\, x_4\, x_3\, x_2\, x_1$

$x(12) \equiv x_{12}\, x_{11}\, x_{10}\, x_9\, x_8\, x_7\, x_6\, x_5\, x_4\, x_3\, x_2\, x_1$

$x(13) \equiv x_{13}\, x_{12}\, x_{11}\, x_{10}\, x_9\, x_8\, x_7\, x_6\, x_5\, x_4\, x_3\, x_2\, x_1$

z is P, L, A, S, or Q; and y is selected from:

$y(0) \equiv$ y is not present

$y(1) \equiv y_1$

$y(2) \equiv y_1\, y_2$

$y(3) \equiv y_1\, y_2\, y_3$

$y(4) \equiv y_1\, y_2\, y_3\, y_4$

$y(5) \equiv y_1\, y_2\, y_3\, y_4\, y_5$

$y(6) \equiv y_1\, y_2\, y_3\, y_4\, y_5\, y_6$

$y(7) \equiv y_1\, y_2\, y_3\, y_4\, y_5\, y_6\, y_7$

$y(8) \equiv y_1\, y_2\, y_3\, y_4\, y_5\, y_6\, y_7\, y_8$

$y(9) \equiv y_1\, y_2\, y_3\, y_4\, y_5\, y_6\, y_7\, y_8\, y_9$

$y(10) \equiv y_1\, y_2\, y_3\, y_4\, y_5\, y_6\, y_7\, y_8\, y_9\, y_{10}$

$y(11) \equiv y_1\, y_2\, y_3\, y_4\, y_5\, y_6\, y_7\, y_8\, y_9\, y_{10}\, y_{11}$

$y(12) \equiv y_1\, y_2\, y_3\, y_4\, y_5\, y_6\, y_7\, y_8\, y_9\, y_{10}\, y_{11}\, y_{12}$

$y(13) \equiv y_1\, y_2\, y_3\, y_4\, y_5\, y_6\, y_7\, y_8\, y_9\, y_{10}\, y_{11}\, y_{12}\, y_{13}$

$y(14) \equiv y_1\, y_2\, y_3\, y_4\, y_5\, y_6\, y_7\, y_8\, y_9\, y_{10}\, y_{11}\, y_{12}\, y_{13}\, y_{14}$

$y(15) \equiv y_1\, y_2\, y_3\, y_4\, y_5\, y_6\, y_7\, y_8\, y_9\, y_{10}\, y_{11}\, y_{12}\, y_{13}\, y_{14}\, y_{15}$

$y(16) \equiv y_1\, y_2\, y_3\, y_4\, y_5\, y_6\, y_7\, y_8\, y_9\, y_{10}\, y_{11}\, y_{12}\, y_{13}\, y_{14}\, y_{15}\, y_{16}$

$y(17) \equiv y_1\, y_2\, y_3\, y_4\, y_5\, y_6\, y_7\, y_8\, y_9\, y_{10}\, y_{11}\, y_{12}\, y_{13}\, y_{14}\, y_{15}\, y_{16}\, y_{17}$

and wherein

x_1 is I, R, M, IQR, V, L, K, F, S, G, Y, SRG, or YQR;

x_2 is H, R, Y, T, S, P, F, N, A, K, G, or V;

x_3 is I, L, M, T, V, E, G, F, or Y;

x_4 is R, S, G, H, A, K, or not present;

x_5 is K, R, I, N, Q, A, IR, RQ, or not present;

x_6 is R, K, S, I, P, Q, E, G, or T;

x_7 is T, K, V, I, A, R, P, or E;

x_8 is N, NV, Y, KI, I, T, DK, H, or K;

x_9 is N, S, K, E, Y, D, I, or Q;

x_{10} is N, Y, S, D, G, or H;

x_{11} is P;

x_{12} is R, I, or K;

x_{13} is T, I, M or A;

y_1 is R, K, Q, G, S, or T;

y_2 is A, V, N, R, K, T, S, F, P, or W;

y_3 is F, I, V, L, W, Y, G, S, or T;

y_4 is Y, V, H, L, F, S, I, T, M, R, VH, or FT;

y_5 is T, A, V, Q, H, I, S, Y, or not present;

y_6 is T, R, I, Q, A, M, or not present;

y_7 is G, E, K, R, T, D, Q, A, H, N, P, or not present;

y_8 is R, Q, E, K, D, N, A, G, S, I, or not present;

y_9 is I, V, R, N, G, or not present;

y_{10} is I, T, V, K, M, R, L, S, E, Q, A, or not present;

y_{11} is G, R, E, K, H, or not present;

y_{12} is D, N, I, R, T, S, or not present;

y_{13} is I, M, ME, L, or not present;

y_{14} is R, G, K, S, E, or not present;

y_{15} is Q, K, or R;

y_{16} is A; and

y_{17} is H, Y, R, or Q.

54. The kit, according to claim 51, wherein said compound(s) are selected from the group consisting of

(a) a-Y-S-N-V-R-N-R-I-H-I-G-P-G-R-A-F-H-T-T-K-R-I-T-b;

(b) a-N-N-N-T-S-R-G-I-R-I-G-P-G-R-A-I-L-A-T-E-R-I-I-b;

(c) a-N-N-N-T-R-K-G-I-F-I-G-P-G-R-N-I-Y-T-T-G-N-I-I-b;

(d) a-N-T-R-K-S-I-R-I-Q-R-G-P-G-R-A-F-V-T-I-G-K-I-G-b;

(e) a-N-N-N-T-R-K-R-(I or V)-T-M-G-P-G-R-V-(Y or W)-Y-(X or T)-(A or T)-G-Q-I-I-b;

(f) a-N-N-N-(I or T)-R-K-(R or S)-I-T-(R or K)-G-P-G-(R or K)-V-I-Y-A-T-G-Q-I-I-b;

(g) a-N-N-N-T-R-K-G-I-Y-V-G-S-G-R-(A or K)-V-T-T-R-(D or H or Q)-K-I-(I or M)-b; and

(h) a-T-R-Q-(R or S)-T-P-I-G-L-G-Q-(A or S)-L-Y-T-T-R-b.

55. An immunological assay for detecting and/or quantifying antibody against HIV in a fluid, said assay utilizing at least one compound, where said compound is not a naturally occurring HIV envelope protein, said compound having the capability of eliciting, and/or binding with, neutralizing antibodies, said compound comprising the principal neutralizing domain, or a segment thereof, of an HIV variant.

56. The immunological assay, according to claim 55, wherein said compound has the formula:

a x G z G y b

wherein x is 0 to 13 amino acids in length;
y is 0 to 17 amino acids in length; and
z is P, A, S, Q, or L; and

either a or b, but not both, may be omitted; either a or b individually may comprise any one of the following: cysteine, a protein or other moiety capable of enhancing immunogenicity, a peptide from an HIV principal neutralizing domain, a peptide capable of stimulating T-cells, or a general immune stimulant.

57. A method for generating broad neutralizing polyclonal or monoclonal antibodies, said method comprising immunization with a mixture comprising compounds, where said compounds are not naturally occurring HIV envelope proteins, said compounds having the capability of eliciting, and/or binding with, neutralizing antibodies, where each of said compounds comprises the principal neutralizing domain, or a segment thereof, of an HIV variant.

58. A method for generating broad neutralizing polyclonal or monoclonal antibodies, said method comprising immunization with a hybrid compound which comprises polypeptides having the capability of eliciting, and/or binding with, neutralizing antibodies, where each of said polypeptides comprises the principal neutralizing domain, or a segment thereof, of an HIV variant.

59. A method for generating broad neutralizing antibodies, said method comprising immunization or an animal or human with at least one compound comprising a principal neutralizing domain, or segment thereof, of an HIV variant followed by immunization of said animal or human with

(a) a mixture comprising compounds, where said compounds are not naturally occurring HIV envelope proteins, said compounds having the capabflity of eliciting, and/or binding with, neutralizing antibodies, said compounds comprising the principal neutralizing domain, or a segment thereof, of an HIV variant; or

(b) a hybrid compound which comprises polypeptides, where said polypeptides are not naturally occurring HIV envelope proteins, said polypeptides having the capability of eliciting, and/or binding with, neutralizing antibodies, said polypeptides comprising the principal neutralizing domain, or a segment thereof, of an HIV variant.

60. A method, according to claim 59, said method comprising immunization with at least one compound having the capability of eliciting, and/or binding with, neutralizing antibodies, said compound having the formula

a x G z G y b

wherein x is 0 to 13 amino acids in length;
y is 0 to 17 amino acids in length; and
z is P, A, S, Q, or L; and

either a or b, but not both, may be omitted; either a or b individually may comprise any one of the following: cysteine, a protein or other moiety capable of enhancing immunogenicity, a peptide from an HIV principal neutralizing domain, a peptide capable of stimulating T-cells, or a general immune stimulant.

61. A method of detecting antibody against HIV in a biological fluid, comprising the steps of:

(a) incubating a compound, other than a naturally occurring HIV envelope protein, said compound having the capability of eliciting, and/or binding with, neutralizing antibodies, said compound comprising the principal neutralizing domain, or a segment thereof, of an HIV variant;

(b) detecting complexes formed between said antibody and said compound.

62. The method, according to claim 61, wherein said compound has the following formula:

x G z G y

wherein x is 0 to 13 amino acids in length;
y is 0 to 17 amino acids in length; and
z is P, A, S, Q, or L.

63. A method for prophylaxis or treatment of HIV infection, said method comprising administering to an animal or human in need of such prophylaxis or treatment a composition comprising an antibody raised against a mixture comprising compounds, where said compounds are not naturally occurring HIV envelope proteins, but have the capability of eliciting, and/or binding with, neutralizing antibodies, each of said compounds comprising the principal neutralizing domain, or a segment thereof, of an HIV variant.

64. The method, according to claim 63, wherein said compounds have the following formula:

a x G z G y b

wherein x is 0 to 13 amino acids in length;
y is 0 to 17 amino acids in length; and
z is P, A, S, Q, or L; and

either a or b, but not both, may be omitted; either a or b individually may comprise any one of the following: cysteine, a protein or other moiety capable of enhancing immunogenicity, a peptide from an HIV principal neutralizing domain, a peptide capable of stimulating T-cells, or a general immune stimulant.

65. A method for stimulating a lymphocyte proliferative response in humans which comprises treating humans in need of stimulation of a lymphocyte proliferative response with at least one compound, said compound is not a naturally occurring HIV envelope protein, said compound having the capability of eliciting, and/or binding with, neutralizing antibodies, said compound comprising the principal neutralizing domain, or a segment thereof, of an HIV variant.

66. The method, according to claim 65, wherein said compound has the formula

a x G z G y b

wherein x is 0 to 13 amino acids in length;
y is 0 to 17 amino acids in length; and
z is P, A, S, Q, or L; and

either a or b, but not both, may be omitted; either a or b individually may comprise any one of the following: cysteine, a protein or other moiety capable of enhancing immunogenicity, a peptide from an HIV principal neutralizing domain, a peptide capable of stimulating T-cells, or a general immune stimulant.

67. A method for treatment of HIV infection, said method comprising administering to an animal or human in need of such treatment a composition comprising one or more compounds wherein said compounds are not naturally occurring HIV envelope proteins, but have the capability of eliciting, and/or binding with, neutralizing antibodies, each of said compounds comprising the principal neutralizing domain, or a segment thereof, of an HIV variant.

68. The method, according to claim 67, wherein said compounds have the formula

a x G z G y b

wherein x is 0 to 13 amino acids in length;
y is 0 to 17 amino acids in length; and
z is P, A, S, Q, or L; and

either a or b, but not both, may be omitted; either a or b individually may comprise any one of the following: cysteine, a protein or other moiety capable of enhancing immunogenicity, a peptide from an HIV principal neutralizing domain, a peptide capable of stimulating T-cells, or a general immune stimulant.

69. A method for assaying a biological fluid for the presence of an HIV variant-specific protein, said method comprising

(a) contacting said fluid with an antibody specific for compounds wherein said compounds are not naturally occurring HIV envelope proteins, but have the capability of eliciting, and/or binding with, neutralizing antibodies, said compounds comprising the principal neutralizing domain, or a segment thereof, of an HIV variant; and

(b) detecting immune complexes as a measure of said variant in said fluid.

70. The method, according to claim 69, wherein said compounds of part (c) have the formula

a x G z G y b

wherein x is 0 to 13 amino acids in length;
y is 0 to 17 amino acids in length; and
z is P, A, S, Q, or L; and

either a or b, but not both, may be omitted; either a or b individually may comprise any one of the following: cysteine, a protein or other moiety capable of enhancing immunogenicity, a peptide from an HIV principal neutralizing domain, a peptide capable of stimulating T-cells, or a general immune stimulant.

71. A method of <u>in vitro</u> lymphocyte stimulation comprising treating lymphoid cells from immune animals with at least one compound, other than naturally occurring HIV envelope proteins, said compound having the capability of eliciting, and/or binding with, neutralizing antibodies, said compound comprising the principal neutralizing domain, or a segment thereof, of an HIV variant.

72. The method, according to claim 71, where said compound has the formula

a x G z G y b

wherein x is 0 to 13 amino acids in length;
y is 0 to 17 amino acids in length; and
z is P, A, S, Q, or L; and

either a or b, but not both, may be omitted; either a or b individually may comprise any one of the following: cysteine, a protein or other moiety capable of enhancing immunogenicity, a peptide from an HIV principal neutralizing domain, a peptide capable of stimulating T-cells, or a general immune stimulant.

73. A method of identifying polyclonal or monoclonal antibodies which are useful in prophylaxis or therapy, said method comprising screening antibody-producing cells for ability to bind with at least one compound, other than naturally occurring HIV envelope proteins, said compound having the

capability of eliciting, and/or binding with, neutralizing antibodies, said compound comprising the principal neutralizing domain, or a segment thereof, of an HIV variant.

74. The method, according to claim 73, wherein said compound has the formula

a x G z G y b

wherein x is 0 to 13 amino acids in length;
y is 0 to 17 amino acids in length; and
z is P, A, S, Q, or L; and

either a or b, but not both, may be omitted; either a or b individually may comprise any one of the following: cysteine, a protein or other moiety capable of enhancing immunogenicity, a peptide from an HIV principal neutralizing domain, a peptide capable of stimulating T-cells, or a general immune stimulant.

75. A synthetic gene which encodes for a polypeptide wherein said polypeptide comprises more than one HIV neutralizing epitope.

76. A synthetic gene which codes for a polypeptide wherein said polypeptide comprises neutralizing epitopes from more than one HIV isolate.

77. The synthetic gene, according to claim 76, wherein said polypeptide comprises neutralizing epitopes from 2 to 20 HIV isolates.

78. The synthetic gene, according to claim 76, wherein one of said neutralizing epitopes is from the HIV isolate designated HIV-MN.

79. The synthetic gene, according to claim 76, wherein said polypeptide comprises neutralizing epitopes from the HIV isolates which have been designated HIV-MN, HIV-IIIB, HIV-RF, HIV-SC, and HIV-WMJ1.

80. The synthetic gene, according to claim 76, wherein the regions encoding neutralizing epitopes from different HIV isolates are separated by amino acid spacers.

81. The synthetic gene, according to claim 80, wherein said amino acid spacers are glycines.

82. The synthetic gene, according to claim 76, wherein said gene codes for a polypeptide having the amino acid sequence shown in Table 13, or an equivalent amino acid sequence.

83. The synthetic gene, according to claim 76, wherein said gene codes for a fusion polypeptide.

84. The synthetic gene, according to claim 83, wherein said gene codes for a polypeptide having the amino acid sequence shown in Table 13A, or an equivalent amino acid sequence.

85. A compound comprising neutralizing epitopes from more than one HIV isolate.

86. The compound, according to claim 85, wherein said compound comprises neutralizing epitopes from 2 to 20 HIV isolates.

87. The compound, according to claim 85, wherein one of said neutralizing epitopes is from the HIV isolate designated HIV-MN.

88. The compound, according to claim 85, wherein said compound comprises neutralizing epitopes from the HIV isolates which have been designated HIV-MN, HIV-IIIB, HIV-RF, HIV-SC, and HIV-WMJ1.

89. The compound, according to claim 85, wherein said neutralizing epitopes are separated by amino acid spacers.

90. The compound, according to claim 89, wherein said amino acid spacers are glycines.

91. The compound, according to claim 85, wherein said compound has the amino acid sequence shown in Table 13, or an equivalent amino acid sequence.

92. The compound, according to claim 85, wherein said compound is a fusion polypeptide.

93. The compound, according to claim 92, wherein said compound has the amino acid sequence shown in Table 13A, or an equivalent amino acid sequence.

94. A prophylactic or therapeutic composition, comprising immune globulin, monoclonal antibodies, and/or polyclonal antibodies generated by immunizing an appropriate animal with a composition comprising a multi-epitope compound wherein said multi-epitope compound comprises neutralizing epitopes from 2 to 20 HIV isolates, wherein one of said isolates is HIV-MN.

95. The composition, according to claim 94, wherein said multi-epitope compound comprises neutralizing epitopes from the HIV isolates designated HIV-MN, HIV-IIIB, HIV-RF, HIV-SC, and HIV-WMJ1.

96. A method for generating broad neutralizing polyclonal or monoclonal antibodies, said method comprising immunizing an appropriate animal with a composition comprising a multi-epitope compound wherein said multi-epitope compound comprises neutralizing epitopes from 2 to 20 HIV isolates, wherein one of said isolates is HIV-MN.

97. The method, according to claim 96, wherein said multi-epitope compound comprises neutralizing epitopes from the HIV isolates which have been designated HIV-MN, HIV-IIIB, HIV-RF, HIV-SC, and HIV-WMJ1.

98. A method for prophylaxis or therapy of HIV infection, said method comprising administering to an animal or human in need of such prophylaxis or therapy a pharmaceutical composition comprising immune globulin, monoclonal antibodies, and/or polyclonal antibodies generated by immunizing an appropriate animal with a composition comprising a multi-epitope compound wherein said compound comprises neutralizing epitopes from 2 to 20 HIV isolates, wherein one of said isolates is HIV-MN.

99. The method, according to claim 98, wherein said multi-epitope compound comprises neutralizing domains from the HIV isolates which have been designated HIV-MN, HIV-IIIB, HIV-RF, HIV-SC, and HIV-WMJ1.

100. A process for stimulating a lymphocyte proliferative response in humans which comprises treating humans in need of stimulation of a lymphocyte proliferative response with a composition comprising a multi-epitope compound wherein said compound comprises a neutralizing domain from 2 to 20 HIV isolates, wherein one of said isolates is HIV-MN.

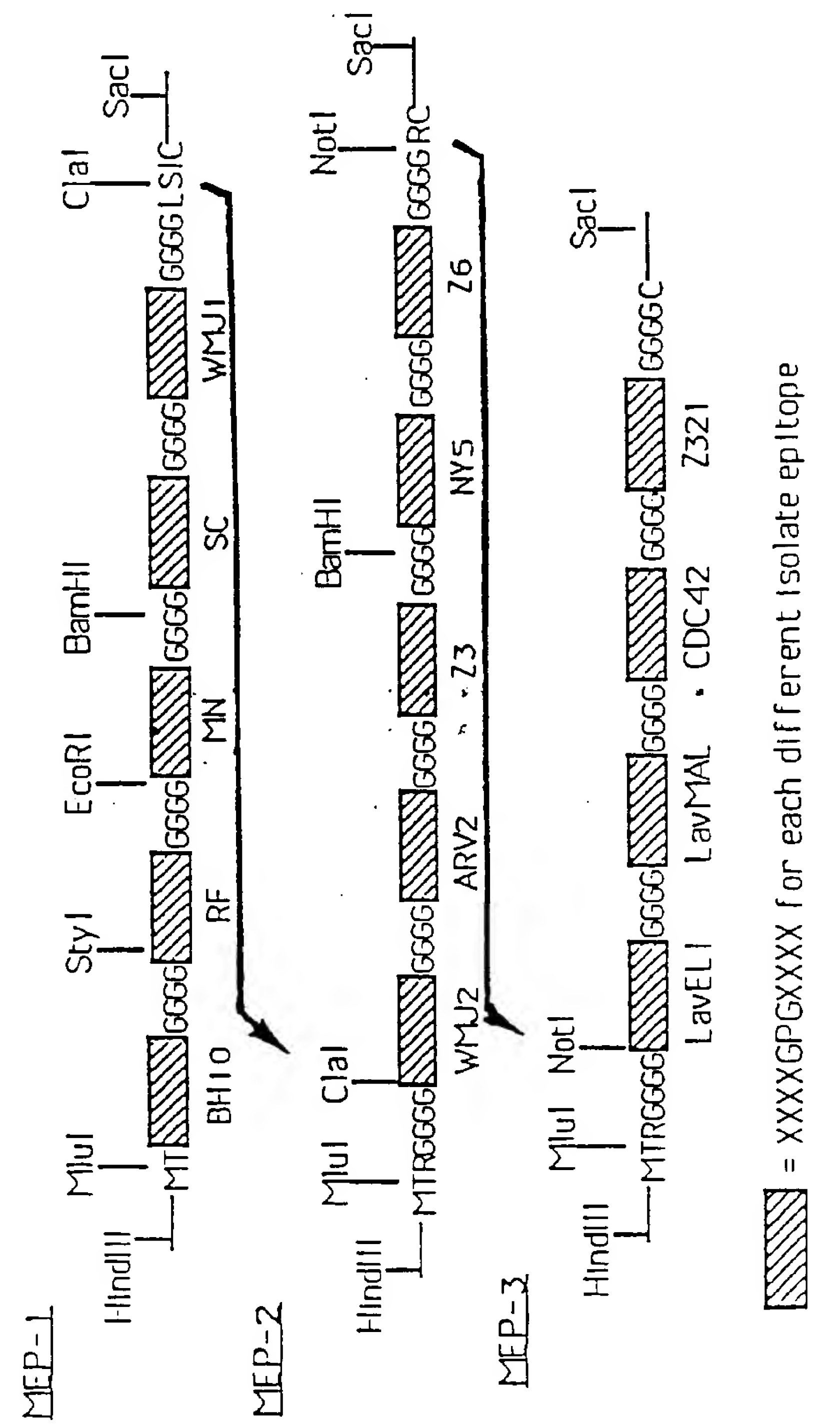
101. The process, according to claim 100, wherein said multi-epitope compound comprises neutralizing epitopes from the HIV isolates which have been designated HIV-MN, HIV-IIIB, HIV-RF, HIV-SC, and HIV-WMJ1.

FIGURE 1

Common Sequence Patterns

OCCURRENCE	K R K R I H I G P G R A F Y T T K
6	V - N - - - - - - - - - - - H - - -
7	T S R G - R - - - - - - - - I L A - E
5	T - - G - F - - - - - - N I - - - G
6	- - Q R - - - - - - - V - I G
4	T - - - $\substack{I\\V}$ T M - - - - - V $\substack{Y\\W}$ - - - G
4	T - - $\substack{R\\S}$ - T K - - - $\substack{R\\K}$ V I - A - G
5	T - - G - Y V - S - - $\substack{A\\K}$ V - - R
4	T - Q $\substack{R\\S}$ T P - - L - Q $\substack{A\\S}$ L - - - $\substack{R\\T}$

FIGURE 2



SUBSTITUTE SHEET

FIGURE 3

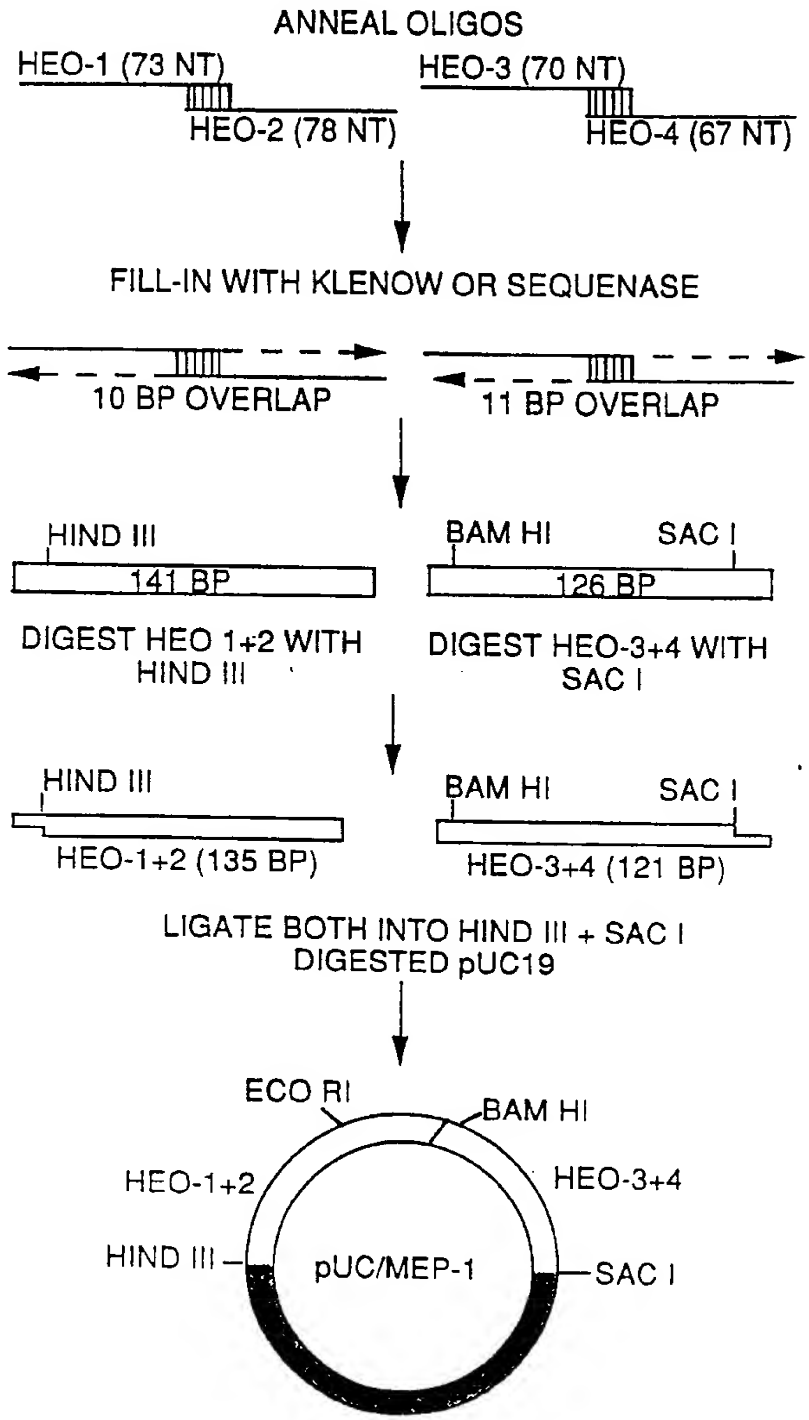

ANNEAL OLIGOS
HEO-1 (73 NT)
HEO-2 (78 NT)
HEO-3 (70 NT)
HEO-4 (67 NT)
FILL-IN WITH KLENOW OR SEQUENASE
10 BP OVERLAP
11 BP OVERLAP
HIND III
141 BP
BAM HI
SAC I
126 BP
DIGEST HEO 1+2 WITH HIND III
DIGEST HEO-3+4 WITH SAC I
HIND III
HEO-1+2 (135 BP)
BAM HI
SAC I
HEO-3+4 (121 BP)
LIGATE BOTH INTO HIND III + SAC I DIGESTED pUC19
ECO RI
BAM HI
HEO-1+2
HEO-3+4
HIND III
pUC/MEP-1
SAC I

FIGURE 4 – OLIGONUCLEOTIDES FOR CONSTRUCTION OF A MULTI-EPITOPE GENE

HEO-1

40
5' CAGTCAAGCT TCCATGACGC GTATCCAGCG TGGTCCGGGT

73
CGTGCTTTTG TTGGTGGCGG AGGCTCCATC ACC 3'

HEO-2

40
5' TAGAAGGCTC TTCCAGGTCC GATGTGAATT CGACCCCCTC

78
CTCCGTAGAT AACCCTTCCT GGTCCCTTGG TGATGGAG 3'

HEO-3

40
5' CGGTGGTGGA GGATCCATAC ATATAGGACC TGGAAGAGCA

70
TTTTATGGTG GAGGTGGTCA CATTCACATC 3'

HEO-4

40
5' GACCGAGCTC AGCAAATCGA TAGGCCCCCT CCGCCGTAGA

67
AAGCACGACC CGGACCGATG TGAATGT 3'

INTERNATIONAL SEARCH REPORT

International Application No PCT/US89/04302

I. CLASSIFICATION OF SUBJECT MATTER (if several classification symbols apply, indicate all) [6]

According to International Patent Classification (IPC) or to both National Classification and IPC

IPC 5: C 07 K 7/10; A 61 K 39/21; C 12 N 15/48; C 12 P 21/02, 21/08; G 01 N 33/569

II. FIELDS SEARCHED

Minimum Documentation Searched [7]

Classification System	Classification Symbols
IPC 5	C 07 K; A 61 K; G 01 N

Documentation Searched other than Minimum Documentation to the Extent that such Documents are Included in the Fields Searched [8]

III. DOCUMENTS CONSIDERED TO BE RELEVANT [9]

Category [*]	Citation of Document, [11] with indication, where appropriate, of the relevant passages [12]	Relevant to Claim No. [13]
X	WO, A1, 88/00471 (SOUTHWEST FOUNDATION FOR BIOMEDICAL RESEARCH) 28 January 1988 see in particular Table II, peptide 2 and pages 20-22 --	1-5,11,12, 19-27,29-32, 34,35,51-62, 69-74
X	WO, A1, 87/02775 (SOUTHWEST FOUNDATION FOR BIOMEDICAL RESEARCH) 7 May 1987 see in particular Table II, peptide 2 and pages 8-11 --	1-5,11,12, 19-27,29-32, 34,35,51-62, 69-74
X	DE, A1, 37 27 703 (GENETIC SYSTEMS CORP.) 5 May 1988	1-62,69-97
Y	see the whole document --	1-62,69-97
X	Journal of Virology, Vol. 62, No. 6, June 1988, p. 2107-2114, S. Matsushita et al: "Characteri-	1-62,69-97
Y	sation of a Human Immunodeficiency Virus Neutralizing Monoclonal Antibody and Mapping of ../...	1-62,69-97

* Special categories of cited documents: [10]

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

IV. CERTIFICATION

Date of the Actual Completion of the International Search	Date of Mailing of this International Search Report
11th December 1989	25 JAN 1990
International Searching Authority	**Signature of Authorized Officer**
EUROPEAN PATENT OFFICE	T.K. WILLIS

III. DOCUMENTS CONSIDERED TO BE RELEVANT (CONTINUED FROM THE SECOND SHEET)

Category *	Citation of Document, with indication, where appropriate, of the relevant passages	Relevant to Claim No
	the Neutralizing Epitope", see the whole document. See in particular fig. 5 and 6 and Table 2 --	
X Y	Proc. Natl. Acad. Sci. USA, Vol. 85, May 1988, p. 3105-3109, H Takahashi et al: "An immunodominant epitope of the human immunodeficiency virus envelope glycoprotein gp160 recognized by class I major histocompatibility complex molecule-restricted murine cytotoxic T lymphocytes", see the whole document and in particular p. 3107, left column and table 2, peptide 18 and p. 3108, last paragraph - page 3109 --	1-62,69-97 1-62,69-97
Y	WO, A1, 87/07616 (BIOGEN N.V.) 17 December 1987 see claim 3, fig. 1, peptide 78, and page 15, lines 13-19 --	7,28,45, 75-97
X	EP, A2, 0 273 716 (THE UNITED STATES OF AMERICA 6 July 1988 see table 3, segment 307-331 and claim 9 --	5,11,12
X	Journal of Virology, Vol. 61, No. 6, June 1987 p. 2024-2028, David D. Ho et al: "Human Immunodeficiency Virus Neutralizing Antibodies Recognize Several Conserved Domains on the Envelope Glycoproteins", see table 2 and page 2027 --	1-4,20-22, 28-32,35
X	Chemical Abstracts, Vol. 109, No. 5, 1 August 1988 (Columbus, Ohio, US) Palker Thomas J. et al: "Type-specific neutralization of the human immu modeficiency virus with antibodies to env-encoded synthetic peptides", see page 470, abstract 36224j, & Proc. Natl. Acad. Sci. U.S.A. 1988, 85(6), 1932-6 (Eng). --	1-5,11,12

ANNEX TO THE INTERNATIONAL SEARCH REPORT ON INTERNATIONAL PATENT APPLICATION NO. PCT/US 89/04302

SA 31698

This annex lists the patent family members relating to the patent documents cited in the above-mentioned international search report. The members are as contained in the European Patent Office EDP file on 08/11/89
The European Patent Office is in no way liable for these particulars which are merely given for the purpose of information.

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO-A- 8800471	28/01/88	AU-D- 77863/87	10/02/88
		EP-A- 0276279	03/08/88
		JP-T- 1500432	16/02/89
WO-A- 8702775	07/05/87	EP-A- 0245362	19/11/87
		JP-T- 63501716	14/07/88
DE-A- 3727703	05/05/88	FR-A- 2603107	26/02/88
		LU-A- 86972	02/03/88
		SE-A- 8703225	21/02/88
		GB-A- 2196634	05/05/88
		NL-A- 8701950	16/03/88
		AU-D- 77201/87	19/05/88
		BE-A- 1000811	11/04/89
		OA-A- 8652	30/11/88
		JP-A- 1085928	30/03/89
		EP-A- 0290893	17/11/88
		AU-D- 15080/88	03/11/88
		JP-A- 1063392	09/03/89
EP-A- 0306219	08/03/89	JP-A- 1095773	13/04/89
EP-A- 0311219	12/04/89	WO-A- 89/03391	20/04/89
		AU-D- 24896/88	02/05/89
		NL-A- 8702403	01/05/89
EP-A- 0302801	08/02/89	WO-A- 89/01516	23/02/89
		FR-A- 2619012	10/02/89
		AU-D- 22691/88	09/03/89
		JP-A- 1157380	20/06/89
WO-A- 8809181	01/12/88	None	
WO-A- 8707616	17/12/87	AU-D- 75404/87	11/01/88
		EP-A- 0269712	08/06/88
		JP-T- 1501547	01/06/89
EP-A- 0273716	06/07/88	WO-A- 88/05051	14/07/88
		AU-D- 13657/88	27/07/88
		JP-T- 63503227	24/11/88

EPO FORM P0479

For more details about this annex : see Official Journal of the European Patent Office, No. 12/82

FURTHER INFORMATION CONTINUED FROM THE SECOND SHEET

V. [X] OBSERVATIONS WHERE CERTAIN CLAIMS WERE FOUND UNSEARCHABLE [1]

This international search report has not been established in respect of certain claims under Article 17(2) (a) for the following reasons:

1. [X] Claim numbers 63-68, 98-101, because they relate to subject matter not required to be searched by this Authority, namely:

See PCT Rule 39.1(iv)

Method for treatment of the human or animal body by means of surgery or therapy, as well as diagnostic methods.

All claims except above mentioned claims and claims 4,8,27,34 and 54 have been searched incompletely,

2. [X] ~~Claim numbers~~ ......, because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

The wordings "neutralizing domain", "neutralizing epitope" and "compounds... are not naturally occuring...." and the broad definition and numerous combinations of the variables x_1, x_2...., y_1, y_2....

a and b do not comply with the requirements of PCT article 6, "claims shall be clear and consise".

3. [ ] Claim numbers........., because they are dependent claims and are not drafted in accordance with the second and third sentences of PCT Rule 6.4(a).

VI. [ ] OBSERVATIONS WHERE UNITY OF INVENTION IS LACKING [2]

This International Searching Authority found multiple inventions in this international application as follows:

1. [ ] As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims of the international application.

2. [ ] As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims of the international application for which fees were paid, specifically claims:

3. [ ] No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claim numbers:

4. [ ] As all searchable claims could be searched without effort justifying an additional fee, the International Searching Authority did not invite payment of any additional fee.

Remark on Protest

[ ] The additional search fees were accompanied by applicant's protest.

[ ] No protest accompanied the payment of additional search fees.

International Application No. PCT/US89/04302

III. DOCUMENTS CONSIDERED TO BE RELEVANT (CONTINUED FROM THE SECOND SHEET)

Category *	Citation of Document, with indication, where appropriate, of the relevant passages	Relevant to Claim No
X	Dialog Information Services, File WPIL, Dialog accession no 4973780, Berzofsky J: "Synthetic peptide corresp. to HIV GP 160 ENV sequence - with elicits sytotoxicity by T cells against HIV and proliferation of HIV-specific cytotoxic T cells", US 7148692, A, 880802, 8837, (Basic) --	5,11,12,15, 65,71,72
P,X	EP, A2, 0 306 219 (REPLIGEN CORPORATION) 8 March 1989 see the claims, example 19 and table 1 --	8,28,45
P,X	EP, A1, 0 311 219 (STICHTING CENTRAAL DIERGENEES-KUNDIG INSTITUUT) 12 April 1989 see in particular column 3, line 41 to column 4, line 11 and examples --	1-5,11-15, 25,26,30-32, 35-43,46-48, 51-53,55,56
P,X	EP, A1, 0 302 801 (INSTITUT PASTEUR) 8 February 1989 see page 14, lines 6-30 and claim 21 --	6,11,17,30, 44
P,X	WO, A2, 88/09181 (TANOX BIOSYSTEMS, INC.) 1 December 1988 see pages 20-24 and claims 69-75 --	1-5,11,12, 19-23,25-32, 34,35,48
A	Nature, vol. 329, 3 September 1987, p. 68, S.E. Adams et al: "The expression of hybrid HIV:Ty virus-line particles in yeast", see the whole document and in particular p. 69, last paragraph - p. 70 --	1,5,9,75-84
A	FEB, vol. 218, no. 2, p. 231-237, M.J.E. Sternberg et al: "Prediction of antigenic determinants and secondary structures of the major AIDS virus proteins", see the whole document. See in particular pages 232-233 "Introduction". --	1-62,69-84
A	Journal of Virology, vol. 61, no. 2, February 1987, p. 570-578, S. Modrow et al: "Computer-assisted analysis of envelope protein sequences of seven human immunodeficiency virus isolates: prediction of antigenic epitopes in conserved and variable regions", see in particular fig. 1, table 2 and page 571 --------	1,5,12